2020-06

Model-based human systems integration

Taranto, Matthew T.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/67549

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

DISSERTATION

MODEL-BASED HUMAN SYSTEMS INTEGRATION

by

Matthew T. Taranto

June 2020

Dissertation Supervisor: Lawrence G. Shattuck

Research for this dissertation was performed at the MOVES Institute.

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** June 2020	**3. REPORT TYPE AND DATES COVERED** Dissertation	
4. TITLE AND SUBTITLE MODEL-BASED HUMAN SYSTEMS INTEGRATION			**5. FUNDING NUMBERS**
6. AUTHOR(S) Matthew T. Taranto			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited. Patents will be pending...withhold release until patent decision is made			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

The practice of Human Systems Integration (HSI) is a multi-domain activity residing at the confluence of several disciplines. HSI currently lacks an accepted unifying theoretical perspective that joins HSI domain resources in terms of total system performance. Systems engineering (SE) has embraced Model-Based SE (MBSE), which signals the need for HSI to consider the development of Model-Based Human Systems Integration (MBHSI). However, HSI is not currently model based. This dissertation examined the efficacy of General Systems Performance Theory (GSPT) and Nonlinear Causal Resource Analysis (NCRA) to model the human system in terms of performance resource capacity and actual performance, execute accurate performance forecasts, articulate the HSI trade space, and address optimization. A laboratory study using a heterogeneous sample measured human basic performance resource capacities (BPR) across 19 cognitive and psychomotor dimensions, then measured novice pilot performance during a simulated Instrument Landing System (ILS) approach in a Cessna-172. Results indicated moderate to strong agreement between predicted and actual performance scores. Additionally, a quantitative approach to articulating HSI trade space and a methodology for facilitating optimization was achieved. This line of research demonstrates MBHSI is a promising approach to improve the capacity of HSI to communicate with containing systems via proven operations research methodologies.

14. SUBJECT TERMS Model-Based Human Systems Integration, Human Systems Integration, General Systems Performance Theory, performance capacity, performance, human performance, total system performance			**15. NUMBER OF PAGES** 297
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

MODEL-BASED HUMAN SYSTEMS INTEGRATION

Matthew T. Taranto
Lieutenant Colonel, United States Air Force
BS, Northern Arizona University, 2003
MS, Embry-Riddle Aeronautical University, 2010
MS, Naval Postgraduate School, 2013
Master of Military Operational Art and Science, Air University, 2017

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN MODELING, VIRTUAL ENVIRONMENTS, AND SIMULATION

from the

NAVAL POSTGRADUATE SCHOOL
June 2020

Approved by: Lawrence G. Shattuck
Department of Operations Research
Dissertation Supervisor

Rudolph P. Darken
Department of Computer Science

Glenn A. Hodges
MOVES Institute

Robert A. Koyak
Department of Operations Research

Eugene P. Paulo
Department of Systems Engineering

Anthony P. Tvaryanas
Federal Aviation Administration

Lawrence G. Shattuck
Department of Operations Research
Dissertation Chair

Approved by: Peter J. Denning
Chair, Department of Computer Science

Orrin D. Moses
Vice Provost of Academic Affairs

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The practice of Human Systems Integration (HSI) is a multi-domain activity residing at the confluence of several disciplines. HSI currently lacks an accepted unifying theoretical perspective that joins HSI domain resources in terms of total system performance. Systems engineering (SE) has embraced Model-Based SE (MBSE), which signals the need for HSI to consider the development of Model-Based Human Systems Integration (MBHSI). However, HSI is not currently model based. This dissertation examined the efficacy of General Systems Performance Theory (GSPT) and Nonlinear Causal Resource Analysis (NCRA) to model the human system in terms of performance resource capacity and actual performance, execute accurate performance forecasts, articulate the HSI trade space, and address optimization. A laboratory study using a heterogeneous sample measured human basic performance resource capacities (BPR) across 19 cognitive and psychomotor dimensions, then measured novice pilot performance during a simulated Instrument Landing System (ILS) approach in a Cessna-172. Results indicated moderate to strong agreement between predicted and actual performance scores. Additionally, a quantitative approach to articulating HSI trade space and a methodology for facilitating optimization was achieved. This line of research demonstrates MBHSI is a promising approach to improve the capacity of HSI to communicate with containing systems via proven operations research methodologies.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

AAF	Adaptive Acquisition Framework
ANAM	Automated Neuropsychological Assessment Metrics
ASVAB	Armed Services Vocational Aptitude Battery
BEP	basic element of performance
BPR	basic performance resource
CDI	course deviation indicator
CI	configuration item
DAG	Defense Acquisition Guide
DAS	Defense Acquisition System
DASDSE	Deputy Assistant Secretary of Defense for Systems Engineering
DME	distance measuring equipment
DOD	Department of Defense
DoDD	Department of Defense Directive
DoDI	Department of Defense Instruction
DoP	dimension of performance
DSOC	Defense Safety Oversight Council
DV	dependent variable
ERM	Elemental Resource Model
FAA	Federal Aviation Administration
FAR/AIM	Federal Aviation Regulations and Aeronautical Information Manual
GPS	Global Positioning System
GSPT	General Systems Performance Theory
HFE	Human Factors Engineering
HLT	high-level task

HLTp	high-level task performance
HSI	Human Systems Integration
INCOSE	International Council of Systems Engineering
ISD	instructional system design
IV	independent variable
KSEA	Seattle-Tacoma International Airport
MBHSI	Model-Based Human Systems Integration
MBHSI-FR	MBHSI functional requirement
MBSE	Model-Based Systems Engineering
MDA	Milestone Decision Authority
MOS	military occupational series
NCRA	Nonlinear Causal Resource Analysis
NDS	National Defense Strategy
NPS	Naval Postgraduate School
NSS	National Security Strategy
PCE	performance capacity envelope
PM	Program Manager
R_A	resource available
RCA	root cause analysis
R_D	resource demand
RDF	resource demand function
RWY	runway
SAB	Scientific Advisory Board
SE	systems engineering
TSP	total systems performance
TTAMS	tool, technique, approach, method, and/or standard
USAF	United States Air Force

USAFSAB	United States Air Force Scientific Advisory Board
VA	Veterans Affairs
VFR	visual flight rules
WRST	Wilcoxon Rank Sum Tests

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The practice of Human Systems Integration (HSI) is a multi-domain activity residing at the confluence of several disciplines, including systems engineering (SE), the Defense Acquisition System, operations research, performance forecasting, economics, human performance, and modeling and simulation. HSI lacks a generally accepted unifying theoretical perspective that joins HSI domain resources in terms of total systems performance (TSP). The warfighter, HSI, SE, and the DOD would benefit from a theoretical perspective that bridges domain considerations with TSP in terms of HSI. SE has embraced Model-Based SE (MBSE), which signals the need for HSI to consider the development of Model-Based Human Systems Integration (MBHSI). However, HSI is not currently model based. This dissertation's mission statement is

> to create an orderly, sensible, theoretically and model-based methodology to develop truly integrated solutions for the warfighter at minimum cost, optimizing the conversion of resources into TSP

This dissertation examined the efficacy of General Systems Performance Theory (GSPT) and Nonlinear Causal Resource Analysis (NCRA) to model the human system in terms of performance resource capacity and actual performance, execute accurate performance forecasts, articulate the HSI trade space, and address optimization. GSPT is defined as "a framework for modeling systems, tasks, and their interface using an abstraction that focuses on performance and all attributes thereof" and is proposed as a unifying theoretical perspective for HSI and MBHSI (Kondraske 2011, p. 238).

The purpose of this new concept, MBHSI, is to improve HSI's capacity to enable SE. The researcher's definition of MBHSI resulted from the synthesis of the DOD HSI objectives, policies, and guidance; definitions offered by INCOSE and Tvaryanas (2010); HSI systemic diagnoses proffered in the past; and HSI's impetus—complexity:

> MBHSI is an essential, model-based, and integrative process that reliably addresses complexity in terms of resource economics while enabling the SE practice. It applies GSPT and NCRA to model and forecast the quantitative relationships between HSI domain resources and system-level performance, targeting the chronic HSI trade space problem and the original objective of

HSI, optimization. Finally, it seeks to communicate its engineering and program management value in engineering terms.

In recognition of HSI's DOD objective, optimization, and in consideration of SE's performance requirement, a decomposition (a form of deduction) walks the reader in reverse from optimization to model inputs. This functional decomposition, what the system of MBHSI should do to satisfy containing system requirements, inductively led to this dissertation's six core MBHSI functional requirements (MBHSI-FRs). The dissertation research study consisted of three iteratively-built research projects (Projects I-III) which investigated, tested, and validated the MBHSI-FRs, producing evidence that MBHSI has merit. Early exploration of GSPT as a theoretical foundation for MBHSI also illuminated the need to recharacterize the HSI domain architecture in terms of GSPT. The purpose of this critical conversion is to provide a constructive bridge from HSI to MBHSI while maintaining the original intent of HSI and adhering to DOD HSI policy (requirements).

A laboratory study using a heterogeneous sample (N = 64) measured human basic performance resource capacities (BPR) across 19 cognitive and psychomotor dimensions, then measured novice pilot performance during a simulated Instrument Landing System (ILS) approach in a Cessna-172. Results indicated moderate to strong agreement between predicted (*Mdn* = 289) and actual performance scores (*Mdn* = 282). A quantitative approach to articulating HSI trade space and a methodology for facilitating optimization are also provided.

This line of research demonstrates MBHSI is a promising approach to improve the capacity of HSI to communicate with containing systems via proven operations research methodologies. This new concept of HSI to datafy (model) the human system in terms of capacity to execute accurate forecasts, quantitative trades, and optimization in a standardized process suggests improved order and meaning for SE and the DOD may be possible.

ACKNOWLEDGMENTS

> This dissertation is dedicated to the American warfighter, past, current, and future. We are forever indebted to your sacrifice, for you live in the arena of consequence. May we always strive to improve your capacity to deliver decisive outcomes for our Nation.

First, I thank my Lord and savior Jesus Christ for all the priceless blessings He has given and entrusted me with. I am truly a blessed man. It was no coincidence that the Bible verse our children were memorizing during the culmination of my dissertation was Philippians 4:13: "I can do all things through Christ who strengthens me."

To my beloved wife, without whom, very little. Thank you for all your sacrifice to make this expedition a reality. Your selfless support is responsible for my success.

To my amazing kids, you are my inspiration and motivation. Thank you for the countless pictures, hugs, and reminders for why I do what I do. I love you both more than you'll ever know. Thank you for enduring this with your mom and me. Our time together in Monterey was a true blessing and you made it so.

To my committee, thank you for holding me to the standard you did. I traveled in the hands of the masters. Your tireless support and mentorship are a reflection of your character and made this work what it is.

Dr. Shattuck, your unrelenting attention to detail and friendly skepticism drove a high standard. Thank you for the sacrifice you made to support this endeavor.

Dr. Tvaryanas, I'll never quite understand your willingness to allow me, of all people, to stand on your shoulders. Thank you for your constant support, insight, feedback, and counsel. This project would have never left the ground without your constant mentorship.

Dr. Koyak, you are the unsung MBHSI hero. Your patience and willingness to translate my ideas into logical code was the true catalyst for MBHSI. Thank you for the countless hours you donated to this cause.

Dr. Paulo, thank you for your constant support, encouraging words, and focus back to systems engineering. This is truly where we make the difference.

Dr. Darken, your curiosity and well delivered questions kept me on my toes! Thank you for your responsiveness and big-picture focus.

Dr. Hodges, thank you for the time you invested in me from the beginning and your vote of confidence. Your straight-shooting, clear-cut style was a great attribute on the team. Your nod to Dr. Gilbert was an early and strong influence in my work. Thank you.

I was also truly blessed to have external support from several people. Dr. Kondraske, thank you for your willingness to support my work and for your efforts years ago in developing General Systems Performance Theory. I hope I did your work justice. Dr. McCauley, thank you for your sage advice, open ear, and the calm between the storms. It was a true gift to spend time with you during my journey. I'll always be grateful for your support. To all of the participants, I thank each and every one of you for your contribution to this work. Your willingness to donate your time was a true gift. Thank you!

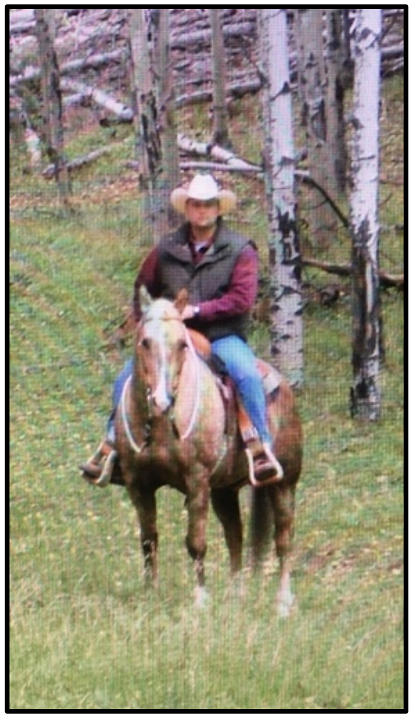
Dr. Dave Cunningham, USAF, Col (Retired), your support of this endeavor allowed this journey to happen. Thank you.

Lt Col Eric Phillips, seriously, thank you for supporting this program from the very beginning. You are a true friend and I am blessed to call you as such.

To my family, thank you all for your constant support and sounding boards. Each one of you played a vital role in my success and I will never forget it.

To my Grandpa "Pots," this one's for you … thank you for encouraging me in my academic pursuits. I knew you were here the whole way.

Finally, I wish to thank Skipper (1996–2019) … you were the truest of friends. I miss you, ol' timer. Come get me at the gate when I get there.



Cloudcroft, NM c. 2005

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

> It must be considered that there is nothing more difficult to carry out nor more doubtful of success nor more dangerous to handle than to initiate a new order of things; for the reformer has enemies in all those who profit by the old order, and only lukewarm defenders in all those who would profit by the new order; this lukewarmness arising partly from the incredulity of mankind who does not truly believe in anything new until they actually have experience of it.
>
> —Nicolo Machiavelli, *The Prince*

A. ORIENTATION

Those who live in the operational environment live in the arena of consequence, and the consequences of failure can be very expensive and extremely dangerous. Unfortunately, there are myriad examples of incidents in which inadequate consideration of Human Systems Integration (HSI) contributed to serious accidents (Booher, 2003; Casey, 1993; O'Connor et al., 2010; Proctor & Van Zandt, 2008; U.S. Air Force, 2012). The Defense Acquisition System (DAS) uses systems engineering (SE) to produce the Department of Defense's (DOD) technology. Therefore, the DAS largely defines this arena via the systems it delivers to the warfighter. SE also serves as the containing system for HSI in the DOD. This means HSI can be described as an enabling system to SE. If the outputs of HSI directly influence SE, then HSI plays an important role in what the DAS delivers to the warfighter. In 2012, the United States Air Force (USAF) Scientific Advisory Board (SAB) found that "over the past 20 years, the capabilities and expertise of the USAF to perform the critical function of HSI have become insufficient" (U.S. Air Force, 2012). A concluding recommendation was to "re-energize the emphasis on Human Systems Integration throughout a weapon system's life cycle, with much greater emphasis during Pre-Milestone A and during Engineering and Manufacturing Development phases" (U.S. Air Force, 2012). What happens in the arena of consequence defines outcomes and their costs. The relationship between them determines value:

$$Value = \frac{Outomes}{Cost.}$$

Figure 1 illustrates an effects-based approach to improve how the DAS delivers value following a revolution in HSI capacity. The first node reflects the purpose of this dissertation, the other nodes identify the series of intended consequences.

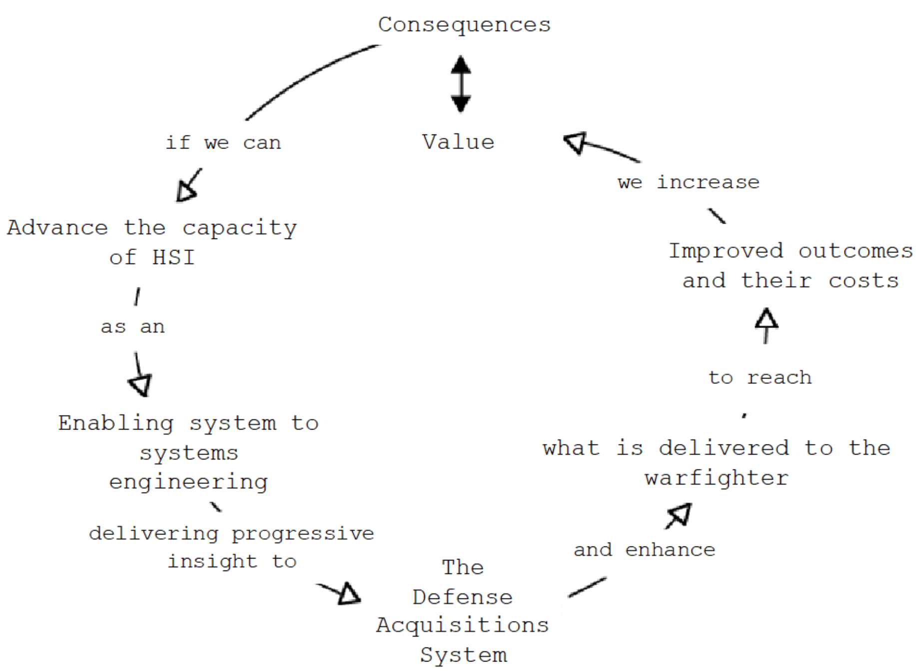


Figure 1. An HSI Effects-Based Approach to Improve System Value in the DOD. Adapted from Hitchens (1992).

Systems rarely generate high value without successfully incorporating humans. While the *Handbook of Human Systems Integration* was published in 2003, it is still widely considered relevant to HSI. When discussing the organizational maturity of HSI, Booher (2003) stated, "Although numerous specific examples of positive human factors influence can be cited, it is fair to conclude that past attempts to incorporate human factors as a primary consideration in government policy for the procurement or regulation of the nation's technology have been marginal at best" (p. 21). Communication issues between

HSI and SE in the DAS likely impede the successful integration of humans into systems. According to Pharmer (2007), "The major challenges with the implementation of HSI have been within acquisition programs themselves because, to some degree, a cultural change has been required to more fully integrate the disciplines of HSI into well-established systems engineering and acquisition activities" (p. 283). In *Putting Systems to Work* (1992), Hitchins offered a diagnosis regarding the HSI symptoms presented by Booher (2003), "that past attempts to incorporate human factors as a primary consideration in government policy for the procurement or regulation of the nation's technology have been marginal at best" (p. 21). Hitchens continues:[1]

> Human factors or human engineering, ergonomics, anthropometrics, etc., have crept into the systems engineering scene, but there still exists something of a gulf between the human factors specialist, focused on the human in his working environment and relating to machinery, and the engineers who design that machinery. They lack a common language; the human factors specialist finds it difficult to be precise in engineering terms about matters of engineering concern, while the design engineer might like nothing better than a transfer function describing a human that he could plug into his calculations. (p. 48)

The gulf Hitchins (1992) speaks of still succinctly depicts the cultural challenges proffered by Pharmer in 2007 and is also consistent with Booher's 2003 appraisal of HSI organizational maturity. If the DAS is to generate high-value systems, then human incorporation must become a higher priority. This re-prioritization requires an improved HSI underpinning, a definition of what optimal means in terms of HSI, and quantitative data to support improved communication between HSI and SE.

The International Council on Systems Engineering (INCOSE) defines SE as "a transdisciplinary and integrative approach to enable the successful realization, use, and retirement of engineered systems, using systems principles and concepts, and scientific, technological, and management methods" (INCOSE, n.d). In describing future trends, Barnes and Beevis (2003) pointed out that "engineering disciplines are responding to the need to maximize system effectiveness, minimize life-cycle costs, and reduce development

[1] While HSI is a relatively new field of practice, older reference material that cite challenges of human factors, human engineering, ergonomics, etc., are still relevant to the present discussion of HSI.

costs and times with a revolution in business affairs" (p. 258). Indeed, the DOD has undertaken something of a revolution by embracing a new approach to SE: Model-Based System Engineering (MBSE) (Office of the Under Secretary of Defense for Acquisition & Sustainment, n.d.). The INCOSE SE Vision 2020 defines MBSE as

> the formalized application of modeling to support system requirements, design, analysis, verification and validation activities beginning in the conceptual design phase and continuing throughout development and later life cycle phases. In particular, MBSE is expected to replace the document-centric approach that has been practiced by systems engineers in the past and to influence the future practice of systems engineering by being fully integrated into the definition of systems engineering processes. (INCOSE, 2007)

The HSI Working Group of INCOSE defined "HSI as the interdisciplinary technical and management processes for integrating human considerations within and across all system elements; an essential *enabler* [emphasis added] to [the] systems engineering practice" (INCOSE, 2007). Booher (2003) claimed that human factors continue to be viewed as supporting elements on unequal "footing with engineering [and] operations disciplines [and that] the challenge for HSI in the twenty-first century is not only to reach an equal footing with these disciplines, but also to actually surpass them" (p. 21). As an *enabling* discipline to SE, HSI needs a complementary model-based approach or risks continued marginalization. Barnes and Beevis (2003) suggested that the expanded use of modeling and simulation must be exploited by HSI or it "will become increasingly difficult for HSI specialists to influence the eventual design solution" (p. 258). In terms of HSI within the DOD, the stated objective is to "*optimize* [emphasis added] total system performance and minimize total ownership costs" (Office of the Under Secretary of Defense for Acquisition & Sustainment, n.d.).

A key point and fundamental tenet of systems thinking suggests that the *containing* systems (DOD and SE) define the function of the *contained* system (HSI). In his 1976 essay titled *Destruction and Creation,* Boyd suggested that in accordance with the Heisenberg indeterminacy principle and the second law of thermodynamics, "we find that uncertainty and disorder generated by an inward-oriented system talking to itself can be offset by going outside and creating a new system" (p. 7). If HSI is contained within SE and the DOD, then HSI must turn to them to define function. Perhaps improved alignment with the HSI-containing systems can improve communication with external systems, thus improving the discipline integration challenges noted by Pharmer (2007). Because the DOD's HSI objective is to optimize performance and minimize cost, optimization requirements define HSI function in part. Additionally, since SE is embracing MBSE, the link now exists between MBSE and Model-Based Human Systems Integration (MBHSI); however, HSI is not currently model based. A simple search of "Model-Based Human Systems Integration" in Google and Google Scholar returned zero matched results, whereas a search of "Model-Based Systems Engineering" returns a rich compilation of sources for the researcher to explore this new model-based-containing system. Therefore, the goal of MBHSI is to support the DOD's optimization goal in the context of SE's focus on performance, which has embraced modeling (MBSE) as a means to achieve total system performance.

Modeling requires establishing relationships between system inputs and outputs. In terms of HSI, inputs are domain considerations including "manpower, personnel, training, human factors engineering, occupational health and safety, force protection and survivability, and habitability" (Department of Defense, n.d.) and outputs are performance data. The Naval Postgraduate School (NPS) HSI model illustrated in Figure 2 provides one conceptual perspective of these relationships in the context of the system acquisition life cycle environment. In terms of science, prediction is a primary goal and theory is a vehicle for achieving this goal (Proctor & Van Zandt, 2008). If theories precede modeling and models are a way to test and refine theories, then modeling HSI will be hard, if not impossible, with the absence of a theoretical perspective that establishes the quantitative relationship between HSI inputs and outputs. Therefore, a theoretical perspective that provides an overarching explanation for how and why one would expect the independent variable(s) to predict the dependent variable is a necessary prerequisite to MBHSI (Creswell & Creswell, 2017). According to Kerlinger and Lee (2000), "A theory is a set of interrelated constructs (concepts), definitions, and propositions that present a systematic view of phenomena by specifying relations among variables, with the purpose of explaining and predicting the phenomena" (p. 11).

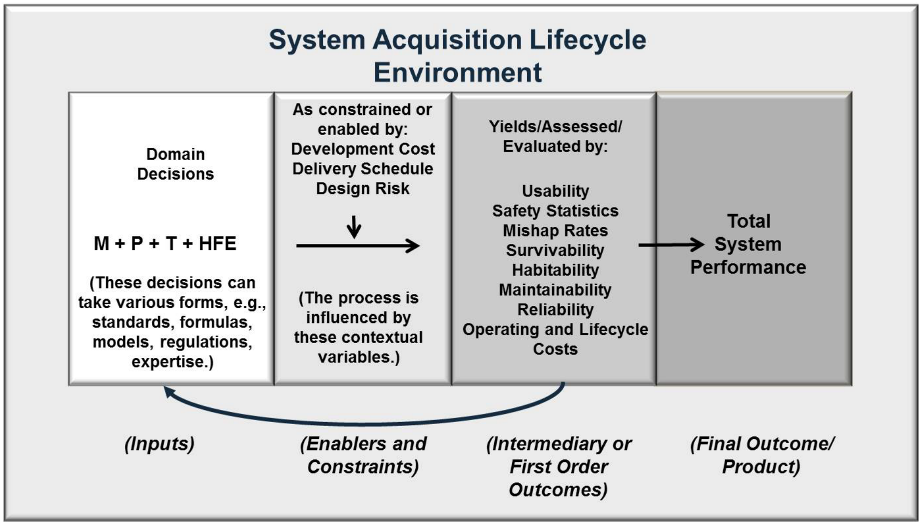


Figure 2. The Naval Postgraduate School HSI Model. Source: Shattuck (2017).

When viewed hierarchically, SE and MBSE both depend on subordinate functions to derive defensible, repeatable, and reliable data using appropriate and accepted methodologies. Fortunately, most of the engineering and science domains not only have accepted theories but also laws (e.g., physics) that provide data for consideration at the higher levels of the engineering hierarchy. Additionally, these domains benefit from the stability of hard systems and subsystems that adhere to mathematical constraints, whereas the human brings a greater amount of uncertainty and complexity. Hard system-resource availability profiles are very similar when compared with like kinds; any two similar models of F-16 aircraft demonstrate nearly identical functional capacities. However, humans display a variety of traits and conditions that lead to a wide range of behaviors; as such, they do not demonstrate identical functional capacities. Hard systems rarely have to contend with such uncertainty and complexity. While specific domains of HSI may have the benefit of theory and the occasional law (e.g., Fitts' law), no unifying theory is generally accepted at the domain integration (or HSI) level. In developing MBHSI as a subordinate system to MBSE, a unifying theory is important to derive defensible, repeatable, and reliable human performance data.

B. THESIS STATEMENT

HSI lacks a generally accepted unifying theoretical perspective that joins HSI domain resources in terms of total systems performance (TSP). The warfighter, HSI, SE, and the DOD would benefit from a theoretical perspective that bridges domain considerations with TSP in terms of HSI. General Systems Performance Theory (GSPT), defined as "a framework for modeling systems, tasks, and their interface using an abstraction that focuses on performance and all attributes thereof" is proposed as a unifying theoretical perspective for HSI *and* MBHSI (Kondraske 2011, p.238). Thus, the consolidated thesis statement is:

GSPT/NCRA can reliably forecast TSP as a function of HSI domain resources.

C. CONTRIBUTIONS

The practice of HSI is a multi-domain activity that resides at the confluence of several disciplines, including SE, DAS, operations research, performance forecasting, economics, human performance, and modeling and simulation. A multi-domain problem space requires a multi-domain solution; this dissertation proposes using GSPT to formulate MBHSI as a complementary approach to MBSE.[2] Reframing HSI in terms of GSPT establishes a new technique in HSI that is focused on enabling MBSE while preserving the historical intent of the practice. MBHSI seeks to interface with MBSE (a containing system) by investigating a theoretically based analytic model that defines the human system in terms of resources and system performance using an evidence-based approach. The objective of this dissertation is to empirically define the relationship between human performance resources and TSP via GSPT and Nonlinear Causal Resource Analysis (NCRA). Specifically, this research examines the efficacy of GSPT/NCRA for modeling the human system in terms of performance resource capacity and system-level performance, executing accurate performance forecasts, demonstrating quantitative system trades, and accomplishing basic system optimization in terms of HSI *and* MBSE. This

[2] Despite the nature of this work, all analytics were completed using Microsoft Excel and R statistical programming language. A provisional patent is being sought by the Naval Postgraduate School Patent Office for the Resource Demand Function (Project II).

dissertation develops a more robust underpinning for HSI to deliver “precision in engineering terms about matters of engineering concern” (Hitchens, 1992, p. 48).

D. MBHSI FUNCTIONAL REQUIREMENTS

According to Boyd (1976), history shows us that if humans “agree to constraints in order to collectively pool skills and talents, obstacles can either be overcome or removed” (p. 1). If HSI is unable or unwilling to identify a robust underpinning (unifying theoretical perspective) that conforms to the containing system (i.e., chooses to go it alone and not agree to constraints upon independent action), obstacles standing in the way of HSI goals may become permanent and continued marginalization is likely. Boyd suggests, “alienated members may dissolve their relationship and form a new group to improve capacity for action” (p. 1).

The overarching thesis statement suggests that the practice of HSI would benefit from a theoretical perspective that reliably ties together the HSI domains in terms of TSP. Such a complementary approach must focus on, and agree to, the requirements of the containing (DOD and SE) systems, such as optimization with a focus on performance. GSPT is proposed as a unifying theoretical perspective for HSI *and* MBHSI due to its capacity to deliver appropriate model-based theoretical outputs to the enabling system while accepting the optimization and performance requirements of the containing systems. From this linear argument, an HSI functional decomposition leads to six core functions to appraise theoretical perspectives and/or tools, techniques, approaches, methods, and standards (TTAMS) for MBHSI. These functions ensure conformity to, and enabling of, the containing systems. They define what MBHSI must do.

In recognition of HSI’s DOD objective, optimization, and in consideration of SE’s performance requirement, this decomposition (a form of deduction) walks the reader in reverse from optimization to model inputs. In terms of optimization, an optimal solution cannot be derived without a known set of feasible solutions. This feasible solution set is defined by quantitative subsystem constraints. These constraints represent measured resources across domains informing system design trade space. Different levels or types of these resources result in a relaxation or restriction of the feasible solution set. Because performance is a requirement and the human system demonstrates intense complexity, performance forecasting is necessary. If complexity management is a requirement of SE, then forecasts must be validated for accuracy. Validation requires actual performance data to quantify the difference between forecasted and observed outputs. Therefore, the relationships between HSI domain resources and performance must be established explicitly and quantitatively. Thus, HSI domain considerations (resources) must be defined in terms of appropriate inputs to the model because the basic principle must be that models drive the data collection and not the other way around (Pidd, 2009). This functional decomposition, what the system of MBHSI should do to satisfy containing system requirements, inductively led to this dissertation's six core MBHSI functional requirements (MBHSI-FRs):

1. Define HSI domain considerations (resources) as model inputs.
2. Establish a quantitative relationship between HSI domain resources and system performance.
3. Define a model that delivers performance data as outputs.
4. Measure performance forecast accuracy.
5. Quantitatively articulate the HSI trade space.
6. Facilitate mathematical program formulation.

The dissertation research study consists of three iteratively-built research projects (Projects I-III). These projects investigate, test, and validate the MBHSI-FRs, producing evidence that MBHSI has merit. If MBHSI is to achieve its purpose and objective, then it should support or satisfy the MBHSI-FRs.

E. DISSERTATION OUTLINE

The first step was to deduce the issues of the DOD HSI objective: optimization. The objective was to pursue a unifying theory that could reliably establish the relationship between HSI-domain resources and TSP. This relationship would result in improved performance forecast accuracy, leading full circle to facilitate optimization in terms of HSI. Boyd's *Destruction and Creation* (1976) served to frame the researcher's thinking and approach to MBHSI. The author's benchmark for success, and hence the desired end state, is a coherently defensible architecture for MBHSI that supports the DOD-stated objective of optimizing system performance and minimizing cost while adhering to the requirements of the containing systems. The purpose is to improve the capacity of HSI to enable SE (Boyd, 1976). MBSE is the current reality; MBHSI is the method to achieve this purpose.

The development of MBHSI is presented in an iterative fashion, a form of synthesis Boyd (1976) called *constructive induction*. This outline provides a blueprint of the structure and content of this research effort. Boyd (1976) suggests that a "crucial step that permits this constructive induction is the separation of the particulars from their previous domains: destructive deduction" (p. 3). This form of "unstructuring is related to deduction, analysis, and differentiation," distilling what MBHSI must do functionally to produce order and meaning for SE (p. 3). Figure 3 illustrates this dissertation's constructive induction of MBHSI.

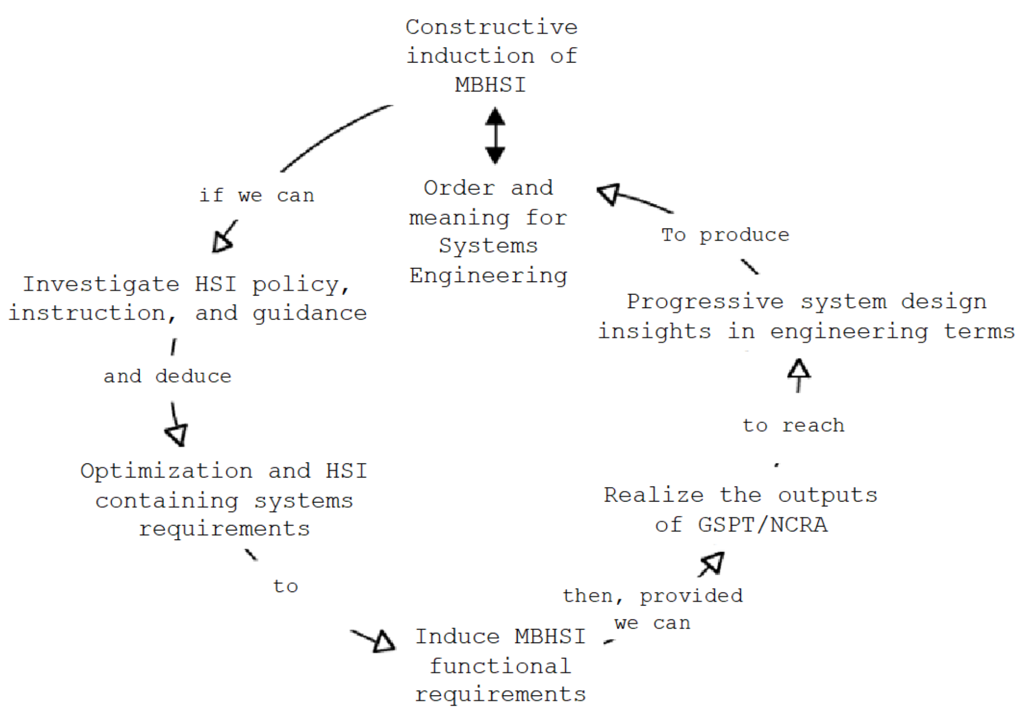


Figure 3. MBHSI Constructive Induction. Adapted from Hitchins (1992).

This dissertation suggests unstructuring and restructuring reveals a way to develop a new concept of HSI. An improved approach to datafy (model) the human system in terms of capacity to execute accurate forecasts, quantitative trades, and optimization in a standardized process may be possible. These desirable outputs of HSI would mark a major increase in HSI capacity and deliver order and meaning for SE and the DOD. Figure 4 develops this dissertation's mission statement, starting with foundation and theory culminating with a future vision. The MBHSI mission statement is

> to create an orderly, sensible, theoretically and model-based methodology to develop truly integrated solutions for the warfighter at minimum cost, optimizing the conversion of resources into TSP.

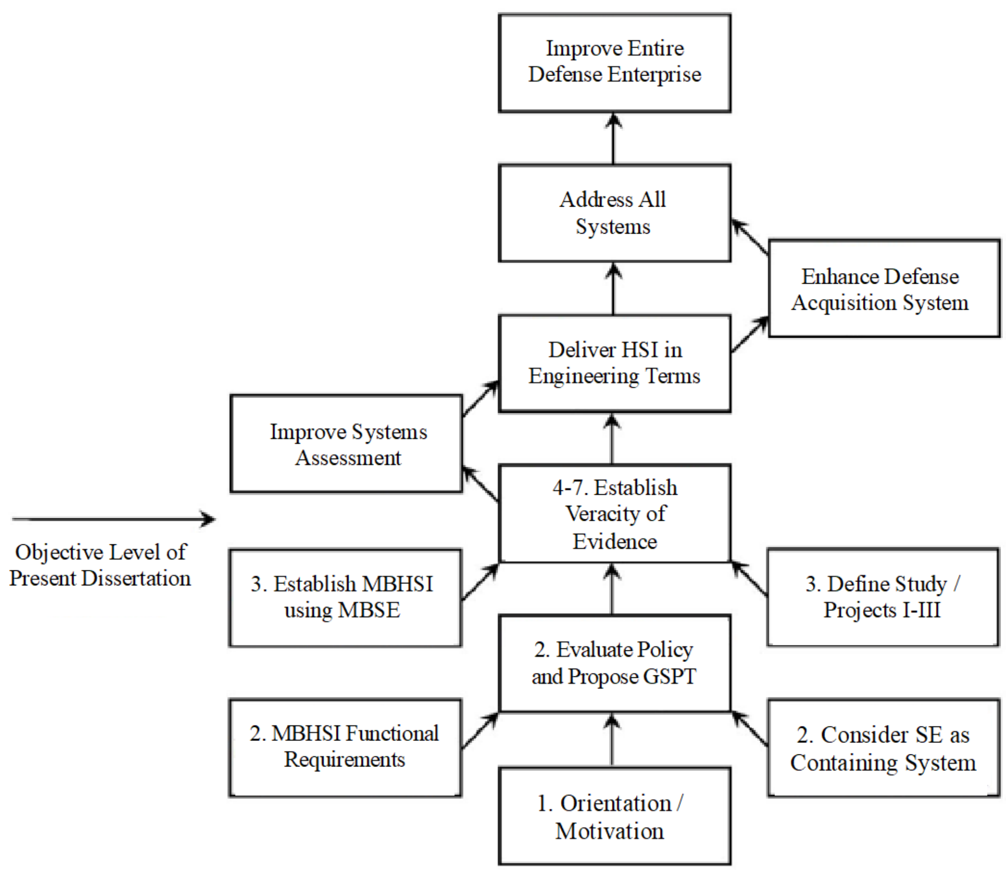


Figure 4. Outlining Structure of Topics and Objectives. Adapted from Tvaryanas (2010).

(1) Chapter II: MBHSI Functional Requirements: Origins and Theory

This chapter outlines HSI policy and DOD objectives to define HSI and its role. SE and MBSE are reviewed, with a focus on how the human system is prioritized and currently integrated in DOD systems design. This chapter includes a review of the 2010 *Defense Safety Oversight Council HSI Task Force* report, GSPT/NCRA, and GSPT literature.

(2) Chapter III: Model-Based Human Systems Integration

This chapter establishes MBHSI using MBSE, SE, and the DOD as reference points leading to a definition of MBHSI. A re-characterization of DOD-recognized HSI domains in terms of GSPT provides a conceptual bridge from HSI to MBHSI. Last, it introduces and defines the dissertation research study as three iteratively-built projects. The projects are mapped to this dissertation's MBHSI-FRs.

(3) Chapters IV-VI: Projects I-III

These chapters present a series of three research projects (introduced in Chapter III) that iteratively develops the overarching thesis statement. It documents evidence in the results sections and highlights unexpected outcomes and threats to validity. The three research projects are:

- Project I: MBHSI Performance Forecast Model Development
- Project II: MBHSI Resource Demand Functions and System Performance Forecasts[3]
- Project III: MBHSI Trade Space

(4) Chapter VII: Addressing a Higher Standard: Optimization

This chapter provides a review of optimization theory, the operations research (OR) process, linear programming (LP), and Dantzig's row method of mathematical program formulation (Dantzig & Thapa, 2006). Chapter VII also articulates the source of variable

[3] A provisional patent will be submitted by the Naval Postgraduate School Patent Office for the MBHSI Resource Demand Function methodology developed for MBHSI.

values in terms of mathematical program formulation using resource and outcome data from Projects I-III. If MBHSI can communicate quantitatively regarding the three fundamental concerns of OR models (decision variables, constraints, and objectives), then truly optimal solutions may be possible in terms of TSP.

(5) Chapter VIII: Discussion

This chapter provides the reader with a review of this research, important findings, and foreshadows future MBHSI work. An appraisal of MBHSI using the MBHSI-FRs demonstrates Boyd's "reversibility and match-up with reality" test (Boyd, 1976). Limitations and threats to validity highlight challenges. A short discussion focused on MBHSI as an approach to HSI that shows potential in enabling MBSE concludes the dissertation.

THIS PAGE INTENTIONALLY LEFT BLANK

II. MBHSI FUNCTIONAL REQUIREMENTS: ORIGINS AND THEORY

Inertia is the first law of history, as it is of physics.

—Morris R. Cohen

A. PURPOSE

This chapter serves two purposes:

1. To examine the origins of the MBHSI functional requirements (MBHSI-FRs) presented in Chapter I through the optimization lens.
2. To support selection of General Systems Performance Theory (GSPT) / Nonlinear Causal Resource Analysis (NCRA) as a potential unifying theory for MBHSI.

Completion of the following objectives demonstrate accomplishment of the purposes:

- a review of current HSI containing systems policy, instruction, and guidance
- a review of SE, identifying the primary requirements that flow to MBHSI (TSP) to include the new DAS Adaptive Acquisition Framework (AAF) and its origins
- a brief introduction to optimization theory as a process for establishing the required variables for mathematical program formulation[4]
- a description of MBSE, its alignment with SE, and how it sets the conditions for MBHSI

[4] Optimization theory highlights the long-standing challenge for HSI to derive values for mathematical program formulation variables.

- a review of the 2010 Defense Safety Oversight Council (DSOC) HSI Task Force recommendation to adopt GSPT as a theoretical approach

- an overview of GSPT/NCRA and relevant literature

Figure 5 illustrates the Chapter II roadmap.

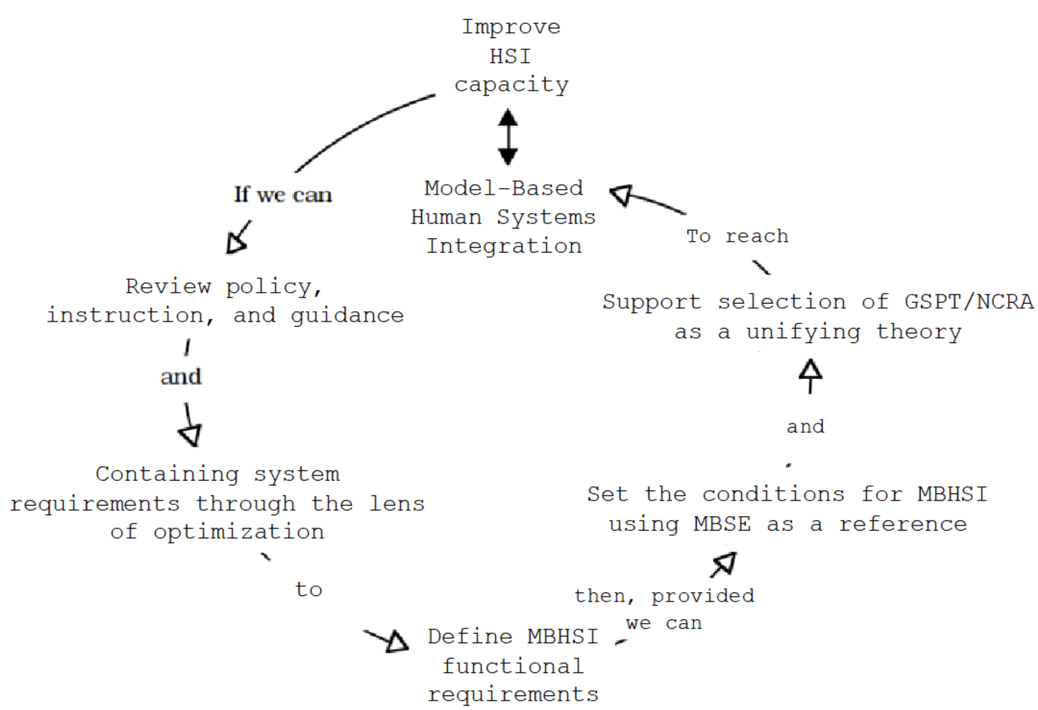


Figure 5. Chapter II Roadmap. Adapted from Hitchins (1992).

B. ORIGINS

1. Background

As mentioned in Chapter I, MBSE is a new approach to the practice of SE. MBSE supports complexity management in part by improving system insights using performance forecasts. Since HSI is not currently model-based, a complementary model-based approach is appropriate. In order to realize a model-based approach to HSI, relationships between system variables (inputs and outputs) must be quantified (Hitchens, 1992), and theoretical perspectives facilitate modeling by explicitly tying together these variables to support forecasting (Creswell et al., 2017). Therefore, MBHSI requires theory to reliably establish model input and output relationships. Figure 6 illustrates the MBHSI system hierarchy as mapped to system requirements.

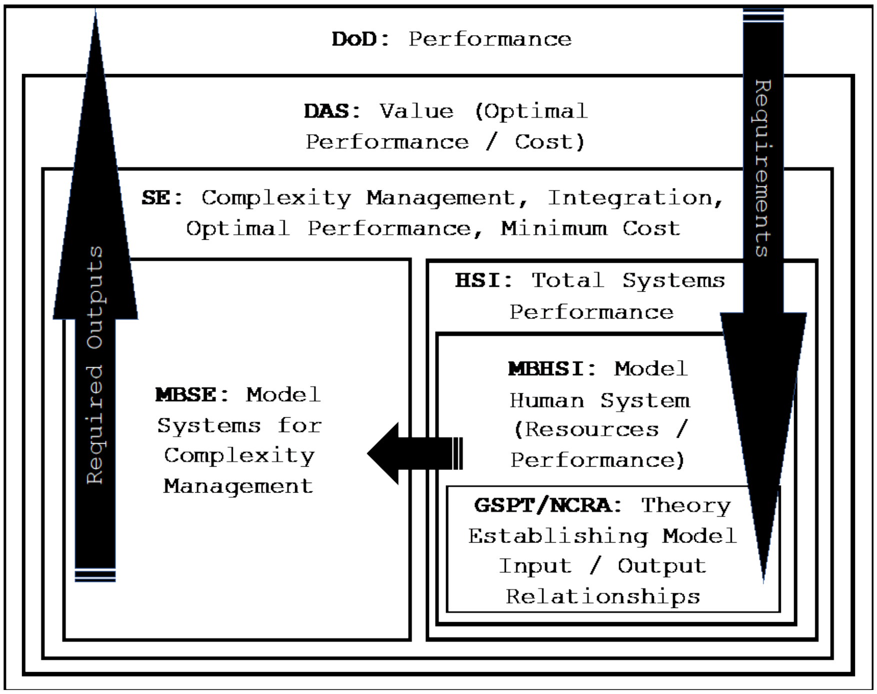


Figure 6. MBHSI System Hierarchy Mapped to Requirements

The practice of HSI involves forecasting future system performance, given system inputs that derive outputs (Tvaryanas, 2010). The DOD's objective for HSI is to optimize TSP, though historically, HSI has been unsuccessful at forecasting TSP in the DOD. To meet this objective, MBHSI must quantify inputs, outputs, and the relationships between the two. Model-based approaches offer promise in improving performance forecasts by quantifying these variables and their relationships; this quantification supports complexity management in 21st-century system design. These realities informed development of the MBHSI-FRs introduced in Chapter I (Figure 7). These functional requirements describe what MBHSI must do as a system to support MBSE and SE. Additionally, any theoretical perspective considered for MBHSI must also support these requirements.

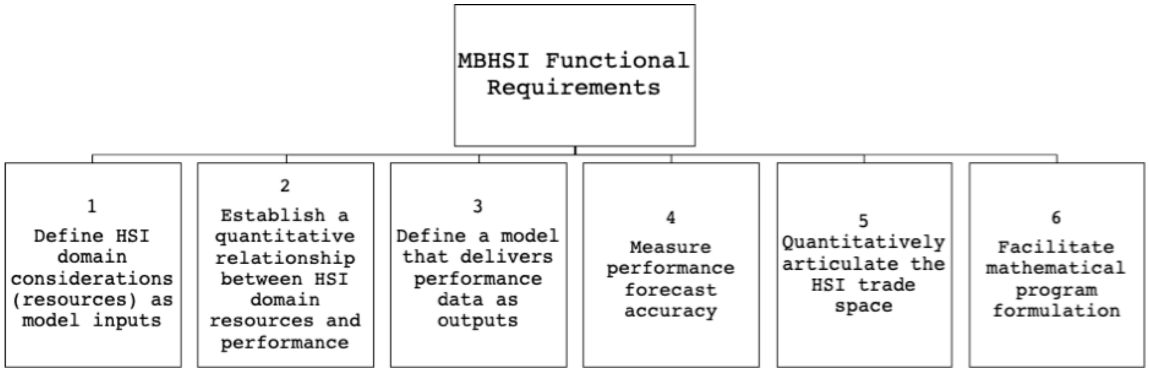


Figure 7. MBHSI Functional Requirements

2. DOD HSI Policy and Guidance

a. Orientation

The MBHSI-FRs reference the following DOD HSI policies and guidance reviewed in this chapter:

- DOD Directive (DoDD) 5000.01: *Defense Acquisition System*
- DOD Instruction (DoDI) 5000.02: *Operation of the Adaptive Acquisition Framework*
    - DoDI 5000.PR (*draft*): *Human Systems Integration in Defense Acquisition*
- The *Defense Acquisition Guide* (DAG)

b. DoDD 5000.01: Defense Acquisition System

The purpose of the DoDD 5000.01 is to "provide management principles and mandatory policies and procedures for managing all acquisition programs" (Department of Defense, 2003, p. 1). The applicability and scope per section 2.1 state, "The policies in this Directive apply to all acquisition programs" (p. 2). Also, the directive states that the DAS is "the management process by which the DOD provides effective, affordable, and timely systems to the users" (p. 2). This definition clearly points to performance (effective) and cost (affordable) requirements from containing systems. The program manager (PM) "is the designated individual with responsibility for and authority to accomplish program objectives for development, production, and sustainment to meet the user's operational needs. The PM shall be accountable for credible *cost* [emphasis added], schedule, and *performance* [emphasis added] reporting to the Milestone Decision Authority (MDA)" (p. 2). Additionally, section 4.1 states, "The DAS exists to manage the nation's investments in technologies, programs, and product support necessary to achieve the National Security Strategy [NSS]" (p. 2). The NSS sets requirements on the development of the *National Defense Strategy* (NDS).

Enclosure 1, section 1.27, "Systems Engineering," states, "Acquisition programs shall be managed through the *application of a systems engineering approach that optimizes total system performance* [emphasis added] and *minimizes total ownership costs* [emphasis added]" (p. 9). This directive establishes the SE requirements of system performance optimization and cost minimization.

The last section in the directive, E1.1.29, "Total Systems Approach," states, "The PM shall apply human systems integration to optimize total system performance (hardware, software, and human) [note the human is last], operational effectiveness [performance], and suitability, survivability, safety, and affordability" (p. 10). Thus, DoDD 5000.01 requires TSP as outputs from HSI to enable SE.

c. DoDI 5000.02: Operation of the Adaptive Acquisition Framework

DOD Instruction 5000.02, *Operation of the Adaptive Acquisition Framework,* is the current instruction from January 23, 2020, that reflects a major restructuring to improve process effectiveness. The policy applicability states, "The DAS supports the National Defense Strategy through the development of a lethal and effective force based on U.S. technological innovation and a *culture of performance* [emphasis added] that yields decisive and sustained U.S. military advantage" (Department of Defense, 2020b). Here, there is an explicit link between the instruction and the NDS: performance.

Historically, Enclosure 7 of DoDI 5000.02 provided specific HSI policy. In March 2020, the researcher obtained a *draft* of DoDI 5000.02PR: "Human Systems Integration in Defense Acquisition." Though at the time of this writing the instruction is in draft form, it is expected to soon replace Enclosure 7. This impending policy served as a reputable source for the stated MBHSI-FRs. Following is a detailed review of this policy to identify critical MBHSI-FRs.

A clear distinction emerges immediately in Section 1.2, "Policy": "It is DOD policy that *HSI* [emphasis added] in defense acquisition *provide a disciplined, unified, and interactive approach* [emphasis added] to integrate human consideration into system design to optimize total system performance and minimize life-cycle costs" (Department of Defense, 2020, p. 1). This requirement directs HSI—not policy—to provide this approach. The MBHSI-FRs reflect a solution-neutral approach addressing this multi-faceted requirement in policy. Additionally, Section 1.2. instructs that "effective, affordable, secure, and supportable solutions for HSI are provided [not to SE, DAS, or DOD; but] *to the user* [emphasis added]" (p. 1). This distinction reinforces the human system priority in TSP design. Each of the six MBHSI-FRs explicitly support this policy requirement by measuring actual human performance resources.

Section 3, "General," points to model-based methods, "Trade-off analyses ensure *human performance data systematically informs* [emphasis added] and facilitates total systems performance" (p. 5). Later, "Planning" lists a minimum of eight management and planning activities required of the PM. Specifically, this section requires the use of modeling and simulation. A model-based approach to HSI addresses this aspect of the new DoDI 5000.02PR. The first MBHSI-FR requires definition of human performance data as model inputs, satisfying this policy requirement. "General," also instructs HSI planning and implementation to focus in part on integrating the seven HSI domains recognized by the DOD. Figure 8 illustrates the DOD-acknowledged HSI domains presented in this policy.

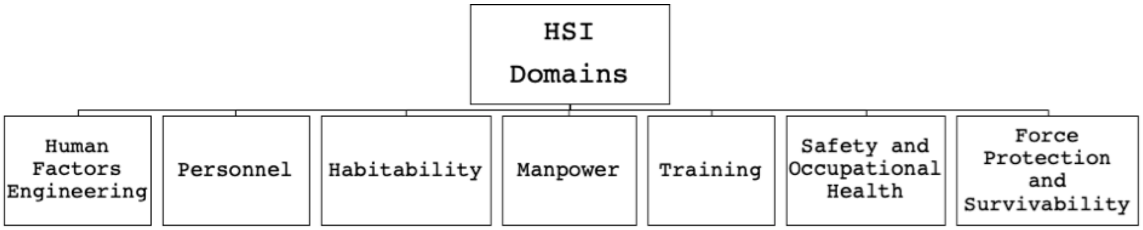


Figure 8. DOD-Acknowledged HSI Domains

Section 3, “Personnel,” instructs the PM to “define human performance characteristics of the user population based on the system description” (p. 6). Additionally, this section directs the PM to “proactively consider personnel availability to ensure operational preparedness” (p. 6). The MBHSI-FRs address this requirement by ensuring measurement of domain resources as inputs. These inputs to the model lead to system performance forecasts. These forecast data, if accurate, define the users in terms of inputs and outputs (performance).

Section 3, “Habitability,” highlights “sustaining system performance” as policy (p. 6). Because the MBHSI-FRs measure performance explicitly, drifts from sustainment levels can be quantified; therefore, this requirement in policy is also addressed by the MBHSI-FRs.

Section 3, “Training,” specifically instructs “individual” training for operators (p. 7). Additionally, this policy requires training decisions to be based on “training effectiveness” (p. 7). This level of effort requires a quantitative understanding of the trainee, the training system, and the system-level performance desired. The MBHSI-FRs address this requirement by measuring resources as inputs to a model, then observing performance outputs. Training can be quantified using a pretest-posttest design that focuses on post-training performance changes. Input and output data then support trades across domains in support of optimal training solutions using the MBHSI-FRs.

Section 3, “Safety and Occupational Health,” requires “integration of safety and occupational health across disciplines and SE” (p. 7). The relationship between the means and ends of this requirement necessitates data that can be integrated. This requirement demands data capable of trading across domains. This MBHSI-FR requires HSI to “execute trades across domains quantitatively" (p. 7). Therefore, the MBHSI-FRs also address this requirement in policy.

Section 3, "Force Protection and Survivability," also clearly prioritizes protection of the human from direct threats and accidents. The policy requires the PM to "reduce susceptibility and probability of personnel being attacked … through system design" (p. 8). The combination of MBHSI-FR performance forecasts, resource inputs, and optimal solution identification seek to address this policy requirement.

In summary, the draft form of DOD 5000.02PR *Human Systems Integration* appeared to pivot significantly toward prioritizing the human in systems design. Despite this pivot, current HSI methods provide great latitude in defining "accomplishment." However, a clear instruction for HSI to "provide a disciplined, unified, and interactive approach to integrate human considerations" squarely places the responsibility on the DOD practice of HSI to deliver that end state (p. 1). As indicated, a model-based approach may prove valuable in achieving the results mandated in this new HSI instruction. Finally, the defined MBHSI-FRs reflect the numerous requirements directed by this impending policy.

d. The Defense Acquisition Guidebook

The DAG claims to provide additional guidance to operationalize and execute HSI in the DOD. Chapter 5 of the DAG "addresses Manpower Planning and HSI in the Defense Acquisition process" (DAU, n.d.-b, p. 1). It provides verbiage such as "including a total systems approach" to incorporate HSI considerations in the process appropriately (p. 1). Chapter 5–3, "Best Practice," recommends "collaboration with the manpower community to examine high-driver tasks" (p. 2). The DAG also recommends a functional analysis to determine which functions to automate, eliminate, etc., to keep the "manpower numbers within constraints" (p. 2). The DAG presents a challenge because without quantifying the relationship between resources at specific levels of performance, forecasting the trade space becomes nearly impossible. Additionally, the DAG recommends manpower requirements be "calculated in conjunction with personnel capabilities, training and human factors engineering trade-offs" (p. 2). Without clear understanding of the desired level of performance by the stakeholder, its achievement may never be realized. Despite this opaque "guidance," the MBHSI-FRs support these stated outputs by using explicit measurement of both resources and performance. Furthermore, the DAG recommends "workload" in determining manpower requirements; however, it does not provide a unit of measure for workload (p. 2). These realities present complex problem sets for the HSI practitioner and dubious insights for SE. Additionally, the DAG focuses on efficiency for cost reasons, yet is devoid of addressing the DOD requirement of performance in calculating manpower for example. If efficiency is sought, then quantitative model inputs and outputs are necessary. It remains unclear how achievement of these HSI "best practices" can be known, albeit positive or negative.

Chapter 5-3.2, "Total Systems Approach," states, "The total systems approach includes equipment, and software as well as people" (p. 3). The human is once again the last consideration, even in the DAG chapter on HSI. Additionally, the DAG states the HSI practitioner is to "assist the PM by focusing attention on the human part of the system" (p. 3). The subjective and qualitative nature of this guidance suggests countless approaches could satisfy this requirement. Later, in Chapter 5-4, "Human Systems Integration," the DAG states, "The key to a successful HSI *strategy* [emphasis added] is comprehensive integration across the HSI domains" (p. 4). The DAG suggested that success depends "on an accurate HSI plan that includes comprehensive integration requirements" (p. 4). It remains unclear how this strategy is derived or what it produces in terms of TSP requirements. However, the MBHSI-FRs support a total systems approach with a distinct focus on the human system, resources, and performance. Therefore, these MBHSI-FRs might provide a feasible strategy for addressing this guidance.

The DAG recommends the domains of HSI "can and should be used to help determine and work the science and technology gaps to address all aspects of the system" without providing guidance to address the relationship between the ends and means of HSI (p. 5). The closest suggestion the DAG provides for a method is, "to accomplish this, HSI domains and human capabilities and constraints should be considered in analytic assumptions and system-of-systems analysis, modeling and testing" (p. 5). This theme is noted throughout the DAG, suggesting a gap between domain inputs and TSP exists.

Considering this, the diagnoses offered earlier by Booher (2003), Hitchens (1992), Madni and Orellana (2018), Pharmer (2007), and Tvaryanas (2010) are accurate.

In summary, the DoDD 5000.01 and DoDI 5000.02 consistently define appropriate requirements for the enabling systems of SE and HSI. Thus, the primary HSI requirement documented in policy is to support TSP. Reviewing these policy documents, there exists a noticeable lack of a unifying construct bridging the gap between HSI domain resources and performance. TSP is a complex problem space, requiring pragmatic solutions to produce value. The MBHSI-FRs provide a reasonable, solution-neutral architecture addressing these requirements. Soft verbs, unclear guidance, and a lack of measurement confirm the unfortunate diagnoses referenced in Chapter I. This evidence further suggests that the practice would benefit greatly from a unifying construct. Establishing a relationship between the inputs and outputs of HSI in terms of appropriate requirements appears to be critical. The next step includes a review of SE, documenting important information regarding systems thinking, the SE process, and the new Adaptive Acquisition Framework (AAF). This review provides context regarding the HSI system hierarchy and further supports the stated MBHSI-FRs.

3. Systems Engineering

"SE is a powerful approach to organizing and conducting complex programs" that continues to evolve to meet the escalating demand signals of complexity, schedules, performance, and efficiency in systems development (INCOSE, n.d.). The origins of SE date back to the 1930s, when British multidisciplinary teams analyzed the air defense system (INCOSE, 2015). Prior to World War II, the "defacto systems engineers were typically architects and civil engineers functioning without the benefit of any defined and consistently-applied processes or practices" (INCOSE Handbook, 2004, p. 9). The RAND Corporation, founded in 1946 by the budding USAF, created *Systems Analysis,* which, according to Buede (2011), "is certainly an important part of SE" because of its focus on outcomes and cost (p. 7).[5] In 1947, Dantzig, a mathematician, was hired by the newly

[5] Note the alignment with the value equation presented in Chapter I, value = outcome/cost.

formed USAF to help solve planning problems. The result was the development of the simplex method, paving the way for linear programming (operations research) and later, the development of the computer (Dantzig & Thapa, 2006). Dantzig's critical contribution of optimization eventually became the DOD objective for HSI, as well as for SE and the DAS (Office of the Under Secretary of Defense for Acquisition & Sustainment, n.d.).[6] The DOD embraced the field of SE in the late 1940s and the missile defense system was the vehicle (Machol & Goody, 1957). An early example of the conceptualization of a system and systems thinking in regard to engineering is noted in the *Air Defense System Engineering Committee* [Valley Committee] Report, October 1950, as cited in Hughes' *Rescuing Prometheus:*

> The world itself is very general … [as for instance] the solar system or the nervous system, in which the work pertains to special arrangements of matter. The Air Defense System has points in common with many of these difference kinds of systems. But it is also a member of a particular category of systems: the category of organisms [defined as] a structure composed of distinct parts so constituted that the functioning of the parts and their relations to another is governed by their relation to the whole. The stress is not only on pattern and arrangement, but on these also as determined by function, an attribute desired in the Air Defense System.
>
> The Air Defense System then is an organism.… What then are organisms? They are of three kinds: animate, organisms which comprise animals and groups of animals, including men; partly animate organisms, which involve animals together with inanimate devices such as in the Air Defense System; and inanimate organisms, such as vending machines. All these organisms possess in common: sensory components, communication facilities, data analyzing devices, centers of judgement, directors of action, and effectors, or executing agencies.… It is the function of an organism … to achieve some defined purpose.

In the 1950s, the space and nuclear races accelerated the complexity and challenges with forecasting emergent attributes of systems, leading to the evolution of SE as a branch of engineering (INCOSE, 2004). Specifically, the Atlas missile project in 1954 was the catalyst (Hughes, 2000). After World War II, an overall sense emerged that system development needed to be interdisciplinary, rather than stove-piped (Hughes, 2000). The

[6] This objective constitutes the requirement of optimization for the enabling system, HSI, by the containing system, the DoD.

creation of a codified, managerial, and engineering science called “systems engineering” to cope with complexity had an early emphasis on science and method. Early SE stressed the putative scientific nature, quantitative approach, and a reliance on theoretical foundations (Hughes, 2000). Thus, the fundamental approach of SE to engage complexity is based on systems thinking, operations research (OR), and systems analysis (performance). Therefore, enabling systems to SE must accept these requirements and deliver outputs to address them—this describes the interface between HSI and SE. Functionally, HSI should deliver outputs in terms of TSP via systems thinking.

In 2002, the international standard ISO/IEC 15288 formally recognized the discipline of SE as a “preferred mechanism to establish agreement between two or more organizations—the supplier(s) and the acquirer(s)” (INCOSE, 2015, p. 15). INCOSE defines SE as “a transdisciplinary and integrative approach to enable the successful realization, use, and retirement of engineered systems, using systems principles and concepts, and scientific, technological, and management methods” (INCOSE, n.d.). Figure 9 depicts the SE “vee” as described by Shishko et al. (1992), to illustrate the iterative process a configuration item (CI) undertakes to pragmatically develop a system. This widely adopted process exists throughout the DAS.

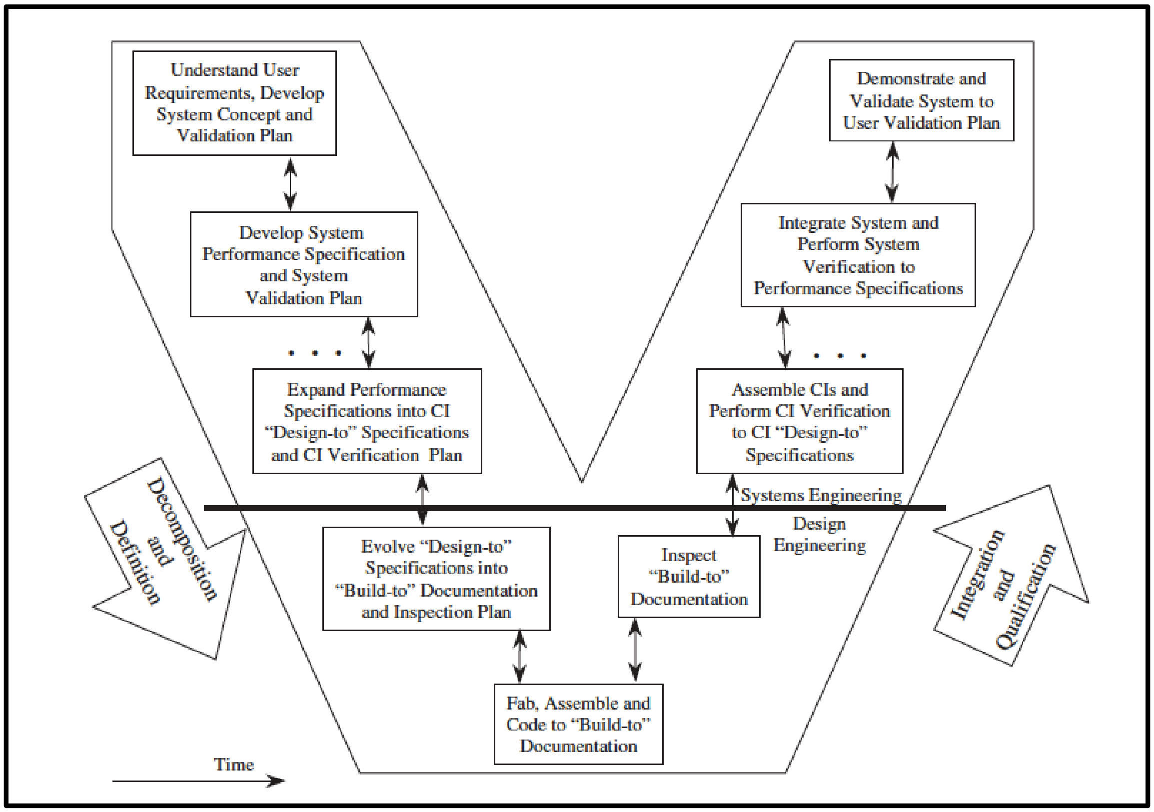


Figure 9. Systems Engineering "Vee." Source: (Shishko et al., 1992).

As a containing system to SE, the DAS recently evolved its process to include "a set of [six] acquisition pathways to enable the workforce to tailor strategies to deliver better solutions faster" (AAF, Figure 10) (DAU, n.d.-a). According to DOD's Undersecretary for Acquisition and Sustainment Ellen M. Lord, "The program will be the most transformational acquisition policy change the department has seen in decades: this policy embraces the delegation of decision-making tailoring program oversight to minimize unnecessary bureaucratic processes and actively managing risks based on the unique characteristics of the capability being acquired" (Garamone, 2019).

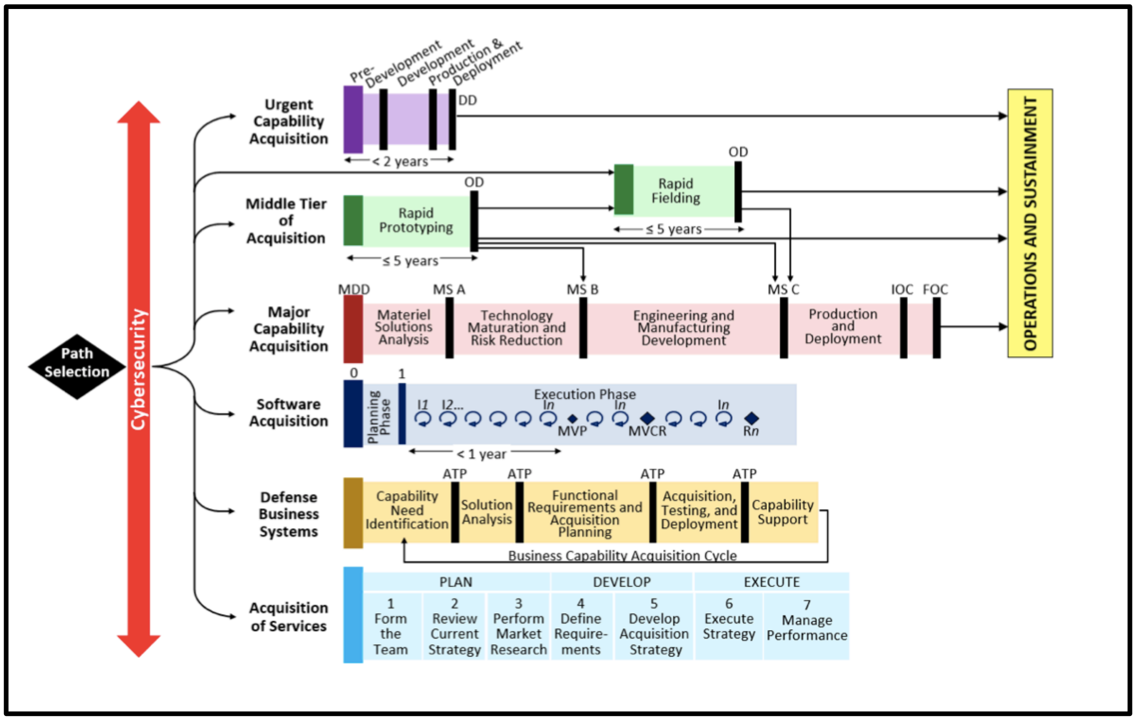


Figure 10. Adaptive Acquisition Framework. Source: DAU (n.d.-a).

This major evolution finds its origins in the 2018 NDS, which defined one of the eleven defense objectives as "continuously delivering *performance with affordability* [emphasis added] and speed as we change Departmental mindset, culture, and management systems" (p. 4). Specifically, the 2018 NDS proffered a call to "Reform the Department for Greater Performance and Affordability" and a "transition to a culture of *performance where results* [emphasis added] and accountability *matter* [emphasis added]" (p. 10). Secretary of Defense James Mattis underpinned this call stating, "Success no longer goes to the country that develops a new technology first, but rather to the one that *better integrates it* [emphasis added] and adapts its way of fighting. Current processes are not responsive to need" (p. 10).[7] This call addresses his strategic approach toward a "culture of performance [that] will generate decisive and sustained U.S. military advantages" by "deliver [ing] performance at the speed of relevance" (p. 10). The requirements of performance and integration therefore must originate from the enabling system through

[7] The explicit call to focus on system integration in the 2018 NDS provides a clear requirement to enabling systems, including HSI.

SE, the DAS, to the DOD. Figure 11 is an example of systems hierarchy from Hitchens (1992, p. 53) that illustrates this critical concept consistent with Boyd's (1976) suggestion that adherence to containing system requirements is critical in defining a new concept and improving capacity for action.

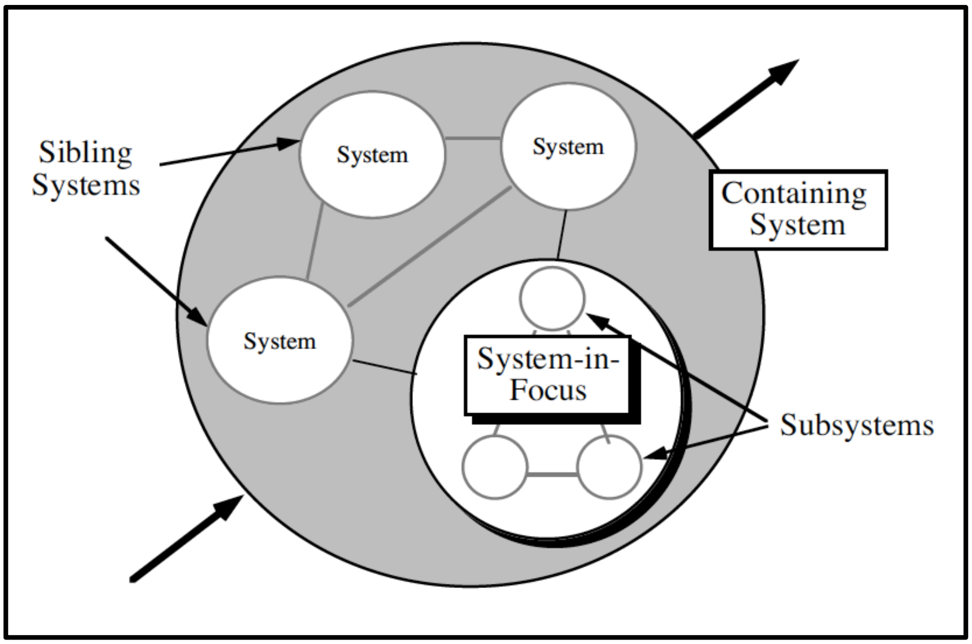


Figure 11. Systems Hierarchy. Source: Hitchens (1992, p. 53).

4. Optimization Theory

Optimization is a detailed process presenting its own set of requirements. According to Dantzig , "Mathematical programming (or optimization theory) is that branch of mathematics dealing with techniques for maximizing or minimizing an objective function subject to linear, nonlinear, and integer constraints on the variables" (Dantzig & Thapa, 2006, p. 1). Formulating linear programs requires abstraction of the problem first, which means a mathematical model must be built before a solution can be found. Dantzig and Thapa (2006) define the mathematical model of a system as "the collection of mathematical relationships which, for the purpose of developing a design characterize the set of feasible solutions for the system" (p. 8). Building mathematical models is as important as solving them because the building “process provides insight about how the system works” and supports information organization (p. 8).

The absence of mathematical modeling in HSI could be explained by Dantzig ‘‘s claim that “models of the real world are not always easy to formulate because of the richness, variety, and ambiguity that exists in the real world or because of our ambiguous understanding of it” (Dantzig & Thapa, 2006, p. 8). HSI lacks a clear method to reliably derive values for the required variables needed formulate mathematical programs. This void prevents desired HSI outputs from transitioning to SE. Establishing a reliable method was a major goal of this dissertation. Projects I-III present novel methods for translating human system data (values) in terms of optimization variables.

Dantzig and Thapa (2006) offers the following system agnostic principles: “The linear programming problem is to determine the values of variables of a system that (a) are nonnegative or satisfy certain bounds, (b) satisfy a system of constraints, and (c) minimize or maximize a form in the variables called an objective” (p. 8). These principles establish the criteria for MBHSI variables in order to achieve the capacity to optimize. Generating appropriate values for these variables in terms of HSI resources was the challenge. The objective function for MBHSI—maximize TSP and minimize cost—addresses (c). MBHSI must establish (b), a system of constraints, and (a) ensures that variables are nonnegative or satisfy certain bounds. Because theory can define relationships between independent and dependent variables, an MBHSI theoretical construct must standardize the establishment of (b) and (a). Reduced ambiguity, mentioned by Dantzig, may be possible as a result of these methods.

Mathematical programing helps identify and organize variables in ways that define feasible solution sets and optimal solutions given decision variables, constraints, and an objective function. Dantzig first describes the column approach and then contrasts it with the row approach (Dantzig & Thapa, 2007). It appears that MBHSI might benefit from using the row approach because of the agreement between limiting resources. Each of Dantzig's steps for the row approach are provided:[8]

1. Define the Decision Variables—"Define all the decision variables that represent the quantity to buy, produce, etc." (p. 11)
2. Define the Item Set—"Determine the classes of objects, the items that are required inputs or are produced as outputs," units of measure required for each type. Choose only those items that are "bottlenecks" (p. 11).
3. Set Up Constraints and the Objective Function—"Write down the constraints associated with the bottleneck by noting how much of each item is used or produced by a unit of each decision variable" (p. 11).

Figure 12 illustrates this row approach to LP formulation. This approach is re-introduced in Chapter VII of this dissertation.

[8] This series of steps in formulating mathematical programs strongly influenced the MBHSI-FRs as it details the requirements to achieve optimization. Recall from Chapter I, the MBHSI-FRs seek to ensure MBHSI satisfies not just containing system requirements, TSP, policy requirements, but also optimization.

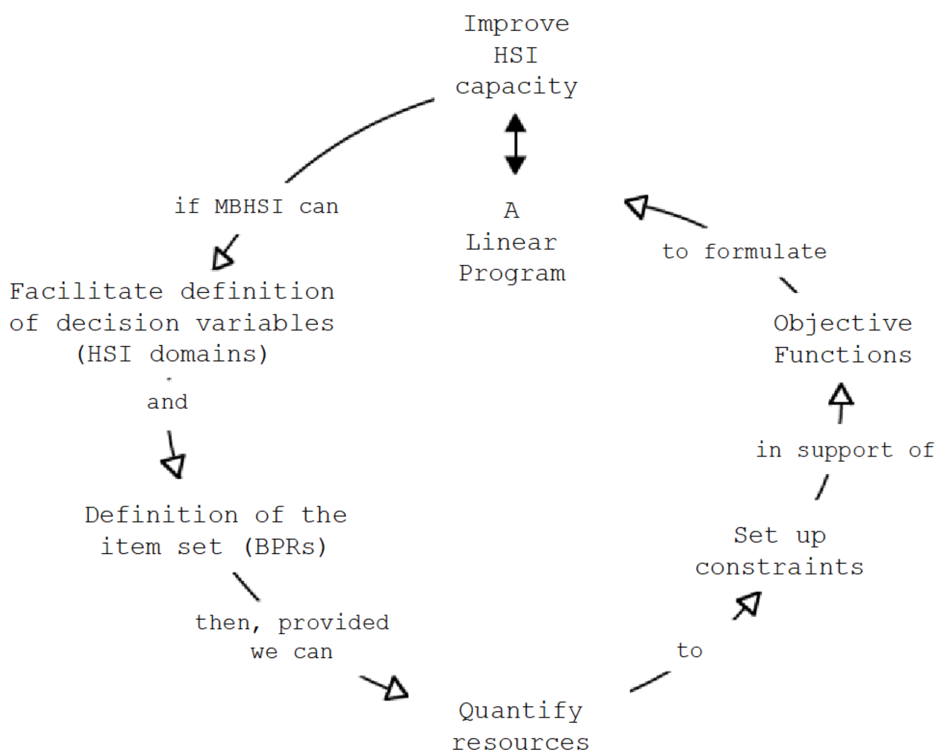


Figure 12. Dantzig's Row Approach to Facilitate Mathematical Program Formulation. Adapted from Hitchins (1992).

Optimization requirements—the primary requirement established by MBHSI-containing systems (TSP) as well as those identified in policy—all influenced the development of the MBHSI-FRs. Therefore, the MBHSI-FRs also define selection criterion for any potential unifying theory for MBHSI. Figure 13 illustrates the concept of hierarchical systems and required outputs.

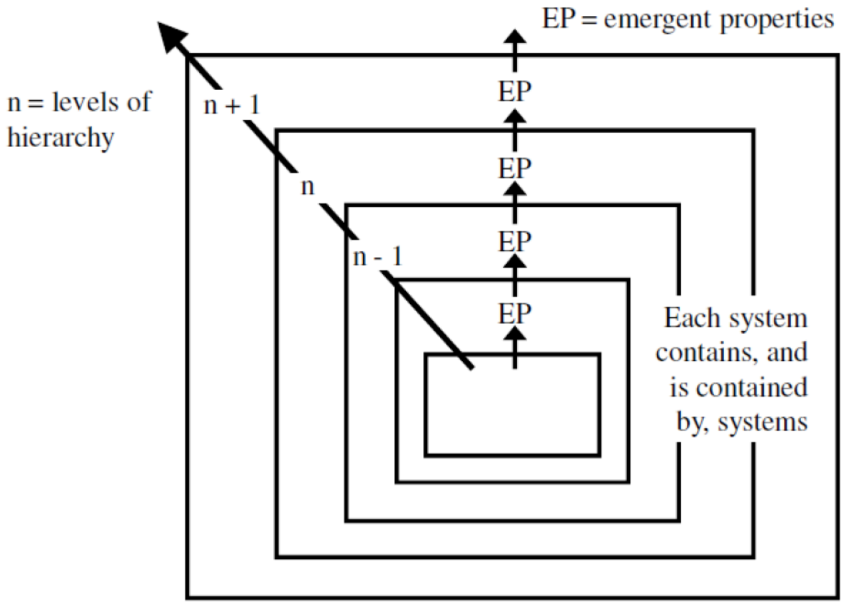


Figure 13. Hierarchy and Emergence. Source: Hitchens (1992, p. 10).

This concept of emergent properties can also be described as the output of an enabling system that achieves some defined purpose (Hughes, 2000). This dissertation refers to this concept as required outputs; those functions required by the containing system: where n is MBHSI, n+1 is MBSE, and n-1 is the MBHSI theoretical concept. Required outputs from theory should enable MBHSI and required outputs from MBHSI should enable MBSE. An introduction to MBSE is provided to establish the reference point for MBHSI.

5. Introduction to Model-Based Systems Engineering

"MBSE is the current paradigm for system engineering" (Madni & Orellana, 2018, p. 1). Recall from Chapter I that the INCOSE SE Vision 2020 defines MBSE as

> the formalized application of modeling to support system requirements, design, analysis, verification and validation activities beginning in the conceptual design phase and continuing throughout development and later life cycle phases. In particular, MBSE is expected to replace the document-centric approach that has been practiced by systems engineers in the past and to influence the future practice of systems engineering by being fully

integrated into the definition of systems engineering processes (INCOSE, 2007).

According to the Office of the Deputy Assistant Secretary of Defense for Systems Engineering (DASD(SE)), "the DOD has realized benefits of using modeling and simulation, but *not* consistently across the acquisition life cycle" (Zimmerman, 2014). Challenges within the DAS presented in Figure 14 include frozen requirements, a linear acquisition process, rigidity, stove-pipes, and limited reuse, among many others. In 2014, DASD(SE) suggested "MBSE as a potential solution to address these challenges" (p. 9). Specifically, MBSE was seen as "part of a long-term trend toward model-centric approaches in SE adopted by other engineering disciplines, including mechanical, electrical, and software" (INCOSE, 2007).

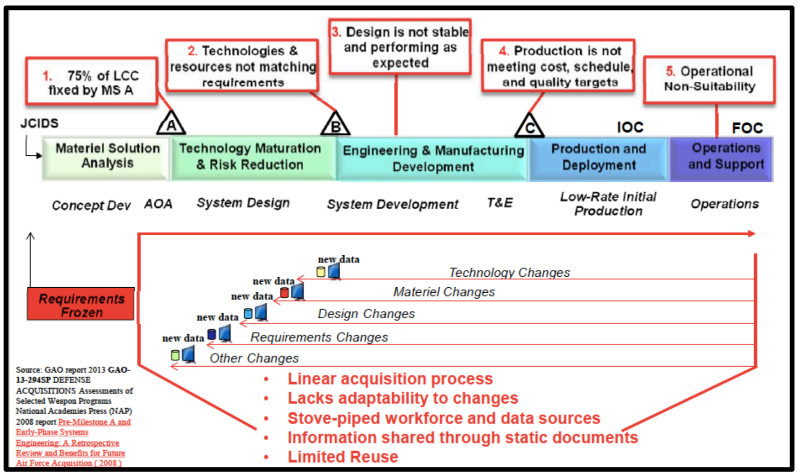


Figure 14. Defense Acquisition System Challenges. Source: Zimmerman (2014, p. 8).

As stated in the INCOSE Vision 2020 Report, "The projected state of MBSE [in] 2020 [was expected to] extend MBSE to modeling domains beyond engineering models to support complex predictive and effects-based modeling [and includes] the integration of engineering models with scientific and phenomenology models, social, economic, and political models and *human behavior models* [emphasis added]" (INCOSE, 2007). These analytic models seek to enable a system-level model. Figure 15 illustrates this MBSE concept.

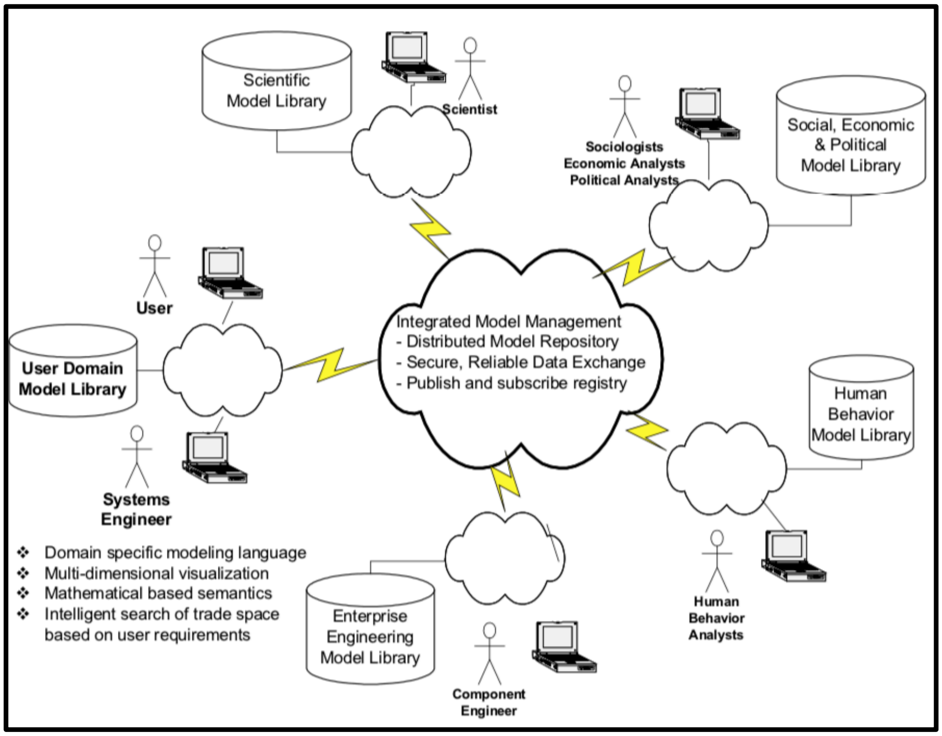


Figure 15. Cross-Domain Model Integration. Source: INCOSE (2007, p. 24).

When fully implemented, MBSE should "reduce the time, cost, and risk to develop, deliver, and sustain systems" (Zimmerman, 2014). Analytic models are thought to "evolve through use in multiple disciplines" and be shareable (Zimmerman, 2014). According to INCOSE Vision 2020, "MBSE is expected to replace the document-centric approach that has been practiced by systems engineers in the past and to influence the future practice of [SE] by being fully integrated into the definition of SE processes" as shown in Figure 16 (INCOSE, 2007, p. 15).

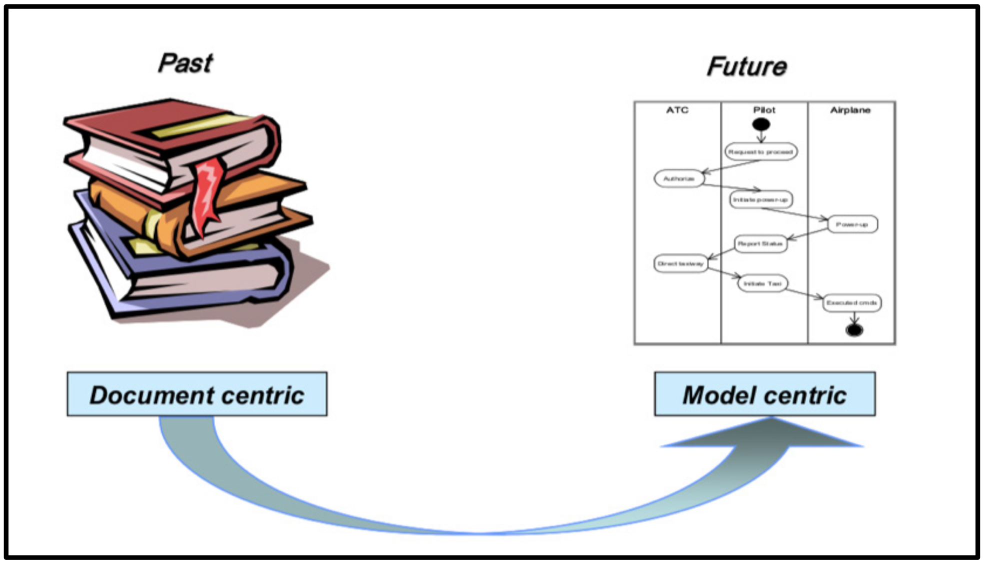


Figure 16. Transition from Document-Centric to Model-Centric Systems Engineering. Source: INCOSE (2007, p. 15).

According to the INCOSE Vision 2025, MBSE design traceability has achieved significant progress. For example, the Ford Motor Company "has established digital design traceability across their onboard electrical and software systems by applying multiple integrated modeling technologies including UML, SysML, and Simulink" (INCOSE, 2007, p. 21). Additionally, when Boeing unveiled the 787 Dreamliner, "it was a virtual rollout. Boeing virtually created parts, and integrated and assembled the system prior to cutting metal" (INCOSE, 2007, p. 21). The identification of "incompatibilities to design interfaces and assembly processes early in design avoid [ed] costly redesign late in the system design life cycle" (INCOSE, 2007, p. 21). However, INCOSE stated, "integration across disciplines, phases of development, and projects represents a key systems engineering challenge" (INCOSE, 2007, p. 22). The document did not mention progress regarding the 2020 projected "human behavior" analytic models.

Madni and Orellana (2018) reported that the human component in systems design has yet to achieve a high-level of priority: "In complex systems, humans are often part of the complex system as opposed to being just users of the system. Compounding the problem is the fact that current system engineering practices tend to address human considerations only as an afterthought" (p. 1).

This appraisal echoes the diagnoses proffered in Booher (2003), Hitchins (1992), Pharmer (2007), and Tvaryanas (2010). However, Madni and Orellana (2018) suggest that this is not an SE burden, stating that "the human factors engineering community has been unable to communicate its value proposition to traditional engineering disciplines and program management" (p. 1). There is strong agreement between Madni and Orellana (2018) and Hitchens (1992), given that Hitchens stated 26 years earlier, "the human factors specialist finds it difficult to be precise in engineering terms about matters of engineering concerns" (p. 48). From that, the author concludes: (a) the contained system must deliver required outputs in terms of the containing system, (b) it is the responsibility of HSI, not SE, to develop and deliver reliable HSI outputs in SE terms, and (c) HSI does not appear to have an answer for (a). This dissertation offers a solution for (c).

Looking forward, the INCOSE SE Vision 2025 "A World in Motion" describes seven "grand challenges" for the SE community (INCOSE, 2014, p. 48). One of those challenges is: "Model-based systems engineering [becomes] a standard practice and is integrated with other modeling and simulation as well as digital enterprise functions" (p. 49). This goal indicates that SE, as a containing system, will continue to pursue MBSE. Additionally, INCOSE's description of the future state of SE infers an additional requirement: "Systems Engineering's theoretical foundations will advance to better deal with complexity … forming the basis for the methods and tools used by practicing systems engineers for architecting, design, and understanding" (p. 24). Specifically,

> methods and tools, based on solid theoretical foundations, will advance to address the demands of innovation, productivity, and time to market as well as product quality and safety by harnessing the power and of advancements in modeling, simulation, and knowledge representation, such as domain-specific vocabularies. (p. 24)

Therefore, if MBHSI is to align with the containing system of SE, a theoretically based approach to MBHSI is appropriate.

INCOSE is focused on "adding fidelity to models, adapting modeling formalisms, and combining multiple concurrent modeling efforts, so systems engineers will be able to perform increasingly detailed trade studies and analyses" (p. 37). Additionally, the council predicts "optimization tools will be used broadly, taking advantage of vast, inexpensive cloud-computing resources to identify system alternatives that are most likely to maximize life cycle value under uncertainty” (p. 37). Given what was achieved through Vision 2020 and what is anticipated through Vision 2025, there appears to be alignment between the concept of MBHSI and the reality of the MBSE trajectory. A callout-box shown in Figure 17 captures INCOSE’s intended translation of current reality to the future and explicitly points to HSI requirements of optimization and performance. This includes an emphasis on value, as presented in Chapter I.

FROM	TO
Systems engineers explore a limited number of design alternatives primarily based on deterministic models of performance, physical constraints, cost and risk.	Systems engineers rapidly explore a broad space of alternatives to maximize overall value, based on a comprehensive set of measures including performance, physical constraints, security, resilience, cost and risk.

Figure 17. Leveraging Information and Analysis for Effective Decision Making. Source: INCOSE (2014).

In the future, SE will be “supported by a more encompassing foundation of theory and sophisticated model-based methods and tools allowing a better understanding of increasingly complex systems and decisions in the face of uncertainty” (INCOSE, 2014, p. 47). This trajectory, along with the continued development of MBSE affirms the need for MBHSI as an enabling system, “ensuring the pieces work together to achieve the objectives of the whole” (INCOSE, 2014, p. ii).

In summary, the author investigated HSI in terms of requirements. These requirements originate in policy, HSI containing systems (SE, DAS, and the DOD), and optimization. This form of functional decomposition, or Boyd’s destruction and creation, defines what HSI must do to achieve its purpose. Each MBHSI-FR can be mapped to requirements in support of enabling HSI’s containing systems. These MBHSI-FRs also define functional criteria for potential MBHSI enabling systems (e.g., theoretical perspectives). Next, a pivot from origins to theory reviews the 2010 DSOC recommendation of GSPT/NCRA for HSI.

C. THEORY

This section examines GSPT/NCRA as a potential theoretical construct for MBHSI. It includes a review of the 2010 DSOC recommendation of GSPT/NCRA, a complete examination of GSPT/NCRA, highlighting key properties and concepts, and a review of GSPT/NCRA in the literature.

1. DSOC Endorsement of General Systems Performance Theory

The DSOC "chartered the Human Systems Integration Task Force to improve safety and reduce mishaps" across the services using HSI (Defense Safety Oversignt Council, 2010, p. 1). According to Major General Thomas Travis (USAF Surgeon General, Ret.), the chair of the Task Force at the time, "in 2009 as part of its findings, the HSI Task Force identified the need for a framework in which to organize human performance" (p. 1). The 2010 DSOC report "represents the dedicated work of twenty-five participants from [18] organizations attacking the problem of finding an organizing framework for human performance" (p. 1). Participants with "expertise in fields touching human performance and HSI from the Army, Navy, Air Force, Marine Corps, NASA, and civilian organizations" contributed to the workshop (p. 1). The workshop quickly "converged on recommending exploration of General Systems Performance Theory and its derivative, Nonlinear Causal Resource Analysis, with potential to address a host of problems faced by the DOD in making the best use of human performance advances in today's dynamic environment" (p. 2). The HSI Task Force concluded:

- There is strong consensus that GSPT and NCRA show great promise as a very capable, evidence-based framework with multiple military applications.
- Correct application of GSPT/NCRA to support decision makers has the potential to: improve personnel selection; reduce attrition; optimize training; inform research in science and technology; improve and validate human centered standards for acquisition requirements, development, and engineering [Human Systems Integration (HSI) standards]; improve performance of individuals, teams, and systems; and thereby, reduce costs and mishaps. (Defense Safety Oversight Council, 2010, p. 1)

Summarized recommendations included:

1. Demonstrate the concept by analyzing existing datasets that will rapidly result in actionable information
2. Fund NCRA analysis of performance data in selected new studies
3. Fund maturation of NCRA software
4. Broaden the understanding of GSPT and NCRA in the safety and human performance communities
5. Follow up with future task force and human performance community meetings to assess and guide progress. (Defense Safety Oversight Council, 2010, p. 3)

Despite these recommendations, no evidence suggests their accomplishment. A decade later, GSPT and NCRA remain a potential framework for HSI; only speculation can speak to the reason(s) these recommendations remain unrealized. This dissertation may be the first pragmatic step in examining GSPT/NCRA for HSI since the 2010 DSOC HSI Task Force recommended it.

2. General Systems Performance Theory and Nonlinear Causal Resource Analysis

a. Overview

In strong consideration of HSI's purpose and following an extensive exploration of theory, methods, and practice spanning a large cross-section of industry, academia, and other domains grappling with complexity, the theoretical construct for MBHSI is GSPT. This systems-agnostic framework models "systems, tasks, and their interface using an abstraction that focuses on performance and all attributes thereof" (Kondraske, 2011, p. 238). "The approach emphasizes a fundamental cause-and-effect, quantitative, systems engineering-oriented performance modeling framework" (Kondraske, 2011, p. 238). GSPT measures system basic performance resources (BPR) in terms of performance capacity. The BPRs define an individual system's N-dimensional performance capacity envelope (PCE). This construct engages resource economic theory, views the human as a system, and measures systems in terms of capacity as noted in Figure 18. Similar to the periodic table of elements in chemistry, GSPT also employs a Monadological (Leibniz et al., 2005) classification system (a finite set of elements for defining more complicated things [i.e., letters of the alphabet]) known as the Elemental Resource Model (ERM) (Kondraske, 2011).

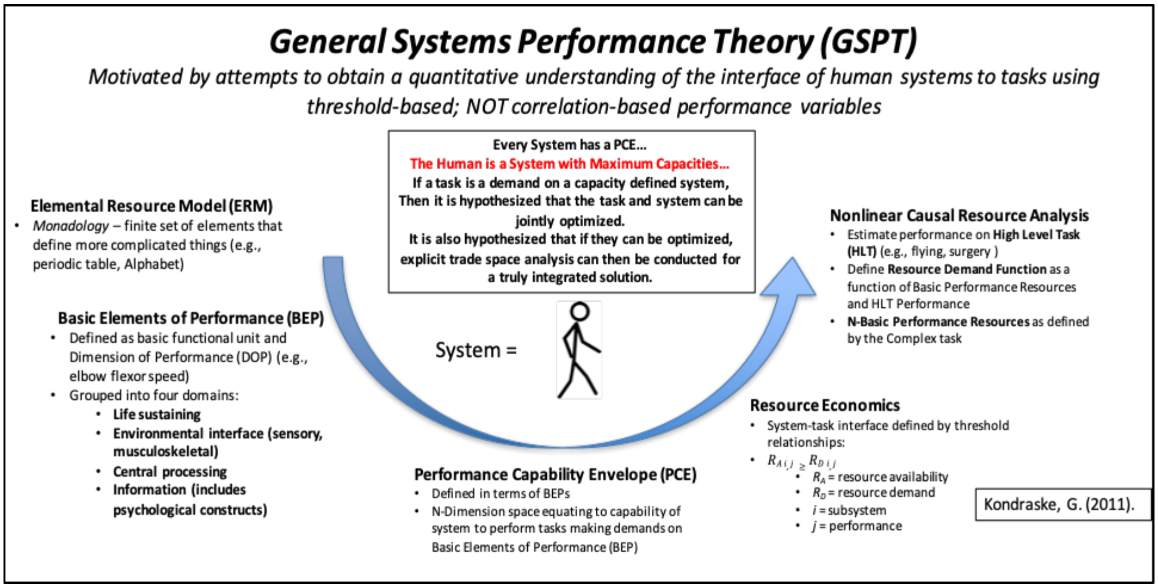


Figure 18. General Systems Performance Theory. Adapted from Kondraske (2011).

The quantitative PCEs for the human support performance forecasts because they define resource availability (R_A). Task performance measurement defines resource demands (R_D). NCRA then establishes the quantitative relationship between the two. This revolutionary approach to HSI forecasts system performance, quantitatively informs trade-offs, and supports optimal solution design. This approach to optimizing system performance for the DOD via improved alignment of supply and demand for human resources is an economic perspective.

b. Key Concepts: GSPT

Selecting GSPT "gives rise to a significant departure in thinking from traditional modeling methods" because it uses a threshold-based approach (Kondraske, 2011, p. 241). Traditional methods "have generally assumed correlation as the expected form of relationship between performance capacities and complex task performance" (p. 241). Human performance resources are the means, GSPT is the way, and high-level task performance (HLTp) is the end. Specifically, GSPT is hypothesized to connect the elusive means and ends of HSI, proposing truly integrated solutions may be possible. Using GSPT/ NCRA to understand and forecast complex system performance, Kondraske (2011) assumes all of the following:

1. "Performance is an omnipresent and ultimate concern." (p. 235)
2. "Performance variables are special and a quantitative understanding of the threshold-based (and not correlation-based) resource economic relationship between performance attributes across hierarchical levels is essential." (p. 235)
3. "Assessment of the notion of change (i.e., in the evolution of complex systems and our understanding of this process) depends, in many situations, first upon valid characterization and understanding of systems from a performance perspective." (p. 235)
4. Any system (to include the human system) can be logically decomposed into a set of performance resources. A performance resource is defined as a functional unit (e.g., finger joint flexor) and a corresponding dimension of performance (DoP) (e.g., strength). A DoP is always defined so that more of a resource is better (e.g., accuracy versus error rate).
5. The system (human)-task interface is defined by resource economics. It is required that $R_{Ai,j} \geq R_{Di,j}$ for all i and j, where "A" represents performance resource availability, "D" represents demand on the given performance resource, i represents the system of interest and j represents a dimension of performance. (p. 241) A task imparts performance resource demands on the human. Human performance R_A must meet or exceed task R_D for all involved resources for successful task performance. Furthermore, the greater the delta between $R_A \geq R_D$, the less risk of task failure. As a result, a threshold (nonlinear) relationship exists between HLTp and lower-level resource availability. R_A will be limiting up to a task-determined threshold value, and thereafter more of the resource will not necessarily result in improved task

performance. This assumption is important as it makes correlational analysis an inappropriate approach to relating human performance resources to HLTp.

6. Monadology: the "idea of using combinations of a set of basic elements or primitives to describe and understand things that are" more complicated (e.g., periodic table of elements, letters of the alphabet). (p. 245)
7. The ERM defines performance at various hierarchical levels of a system as illustrated in Figure 19:
   - **Basic level:** Basic Elements of Performance (BEP) of which visual information processor speed is an example. The human system is comprised of a finite set of BEPs that could be cataloged into a Table of Human Basic Elements of Performance, and in turn, organized into specific series or domains:
     - Life sustaining elements
     - Environmental interface elements (i.e., sensory and biomechanical BEPs)
     - Central processing elements (i.e., neurocognitive BEPs, often referred to as "abilities")
     - Information elements (i.e., learned knowledge and skills as well as psychological BEPs such as personality and motivation)

- **Intermediate level:** Intermediate-level performance resources are derived from sets of BEPs. For example, hand grip strength derives from the synthesis of multiple finger joint flexors' strength.
- **Higher level:** Higher-level performance resources are derived from the synthesis of multiple intermediate level performance resources. They determine a human's ability to perform a set of complex tasks that comprise a job (e.g., RPA pilot).

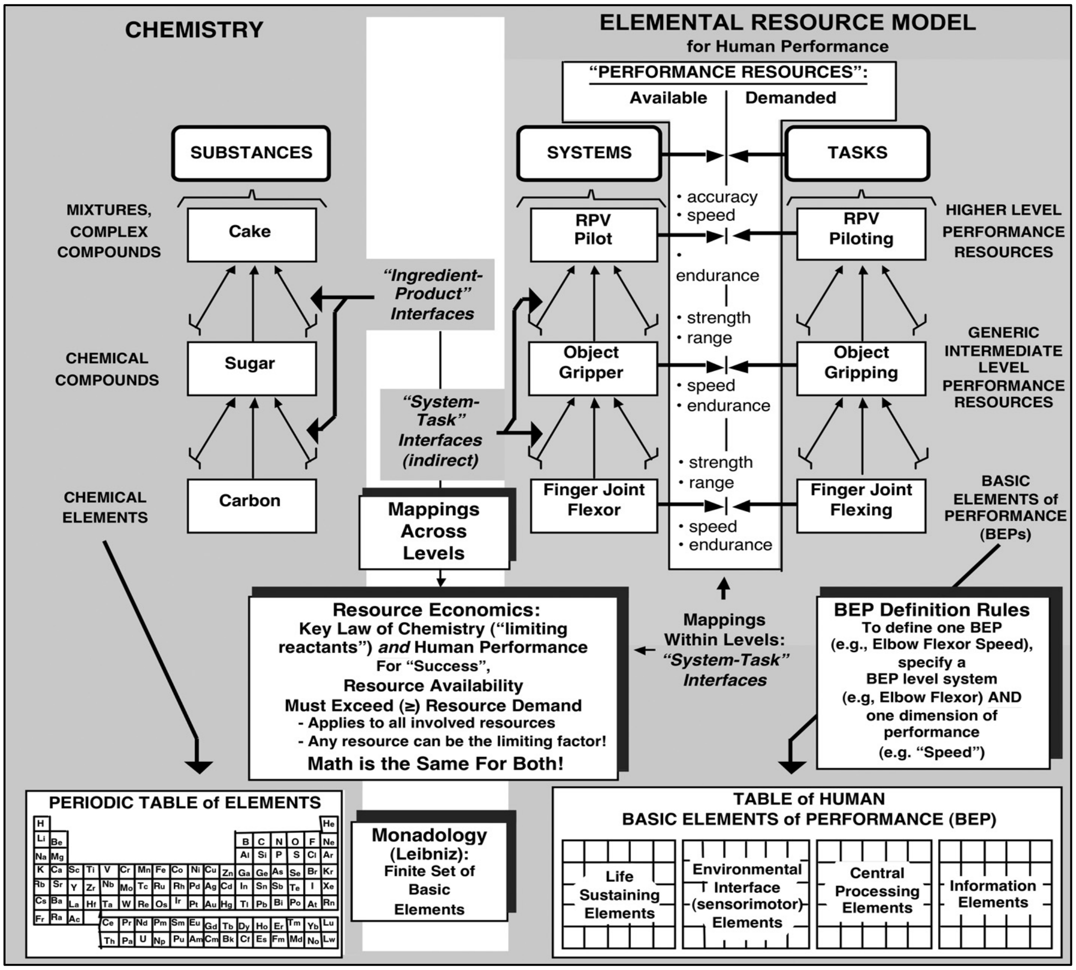


Figure 19. Elemental Resource Model. Source: Kondraske (2011, p. 246).

8. "Modern chemistry and the ERM for human performance share key aspects. Both utilize the notion of hierarchical levels. Performance resources at the BEPs are finite in number, as dictated by the finite set of human subsystems and respective DoPs. BEPs represent the fundamental building blocks of all HLTs. The BEPs are analogous to chemical elements in the periodic table, which are drawn upon to realize all substances (compounds, etc.). GSPT constructs are applied at any hierarchical level, resulting in the identification of the system, its function, DoPs, performance resource availabilities (system attributes), and performance resource demands (task attributes)." (p. 245)

9. With GSPT, a key question is, "How much of a given resource is required to achieve a pre-determined level of HLTp" (Kondraske, 2011, p. 251)? During GSPT work, the HLTp encountered is the known quantity, which is derived from subjective, Gestalt methods in complex tasks.

10. "NCRA provides a means to estimate the degree of performance in an HLT supported by a set of lower-level or more BPRs and also identifies which BPR limits HLTp for a given system (e.g., person, in human system contexts). Of interest is the lower boundary of points in such plots, representing what is called a Resource Demand Function (RDF) in NCRA." "The RDF relates the amount of BPR required to achieve a given degree of HLTp. In this sense, a RDF represents the result of a task analysis" (p. 249).[9] The RDF forecasting the lowest HLTp is the forecast for system-level performance. See Figures 20 and 21.

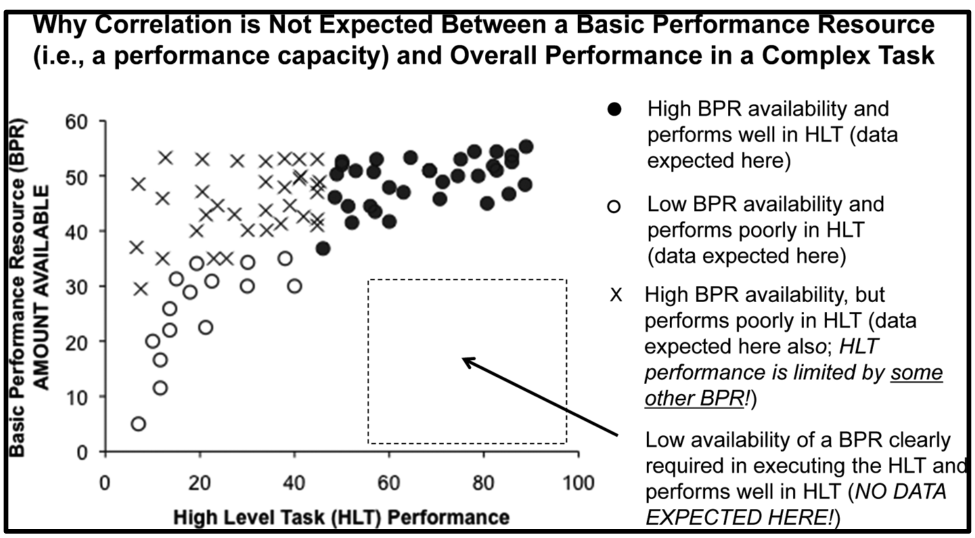


Figure 20. Correlation vs. Lower-Level Performance Variables. Source: Kondraske (2011, p. 249).

---

[9] A significant point of emphasis is noted where in NCRA, the task analysis is *not* an input but rather it is an output.

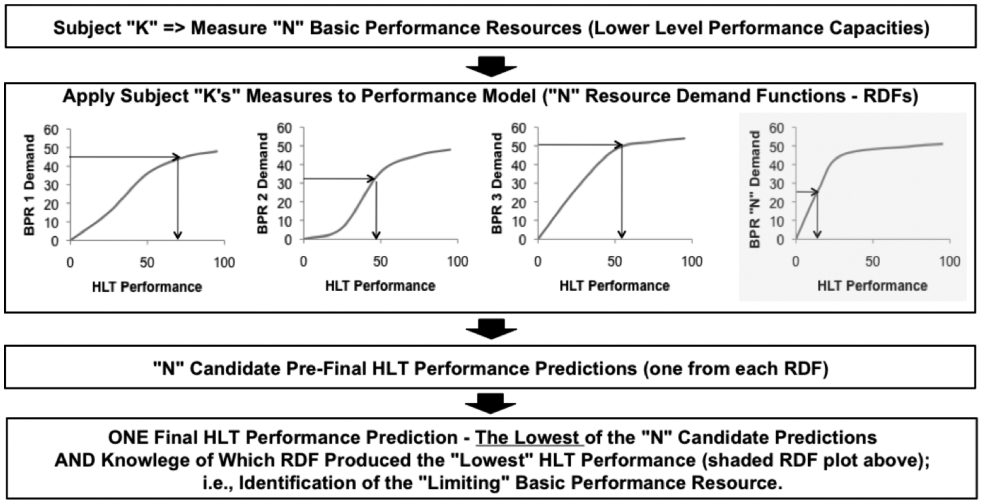


Figure 21. Summary of NCRA Forecast Process.
Source: Kondraske (2011, p. 249).

In summary, two methods "develop and manipulate mental concepts to represent observed reality" (Boyd, 1976, p. 2). The first is to "start from a comprehensive whole and break it down to its particulars or we can start with the particulars and build towards a comprehensive whole" (p. 2). The first method is related to analysis and differentiation and the second corresponds to synthesis and integration (Boyd, 1976).[10] The ERM provides a defensible construct to proceed via deduction or induction, to go from general to specific or specific to general in search of limiting reactants (BPRs) defining product yield, HLTp. Specific economic theory key concepts are employed in both GSPT and MBHSI relevant to performance forecast model understanding and development.

c. GSPT/NCRA: Some Key Distinctions

Initially in GSPT, performance data define the independent variables and BPRs define the dependent variables, a critical distinction; therefore, a logical examination is warranted. According to Dr. Kondraske (2011), the father of GSPT/NCRA:

> Some may question the assignment of independent (High Level Task performance) and dependent (a lower level performance capacity) variables

[10] Note: Not only were Boyd's methods in "Destruction and Creation" used to engineer MBHSI, they were explicitly used during the execution of MBHSI. This evidence supports the notion that MBHSI passes his "reversibility and match-up with reality" test.

> since HLT performance can be viewed to depend on lower level subsystem capacities. However, during GSPT development HLT performance [is] encountered as the known quantity (often measured with subjective, Gestalt methods in complex tasks). A *key question of interest* [is], "*How much of a given lower level performance resource* [*is*] *required (or utilized) to obtain a particular degree of HLT performance?*" This *unknown quantity* [is] viewed to be a function of HLT performance; i.e., a greater amount would be required at high degrees of HLT performance [HLTp] (thus the name "resource demand function"). (p. 251)

This is further discussed elsewhere (Kondraske, 2009) with relevance to interesting history in economics (Gordon, 1982; Lipsey & Chrystal, 2007).

The overarching conditional expectation is simply the expected value of resource given performance, or more succinctly: E [Resource | Performance]. Following development of GSPT models, a reverse conditional expectation achieves considerable value: E [Performance | Resource]. Chapters IV and V demonstrate these explicitly. Figure 22 illustrates this critical concept.

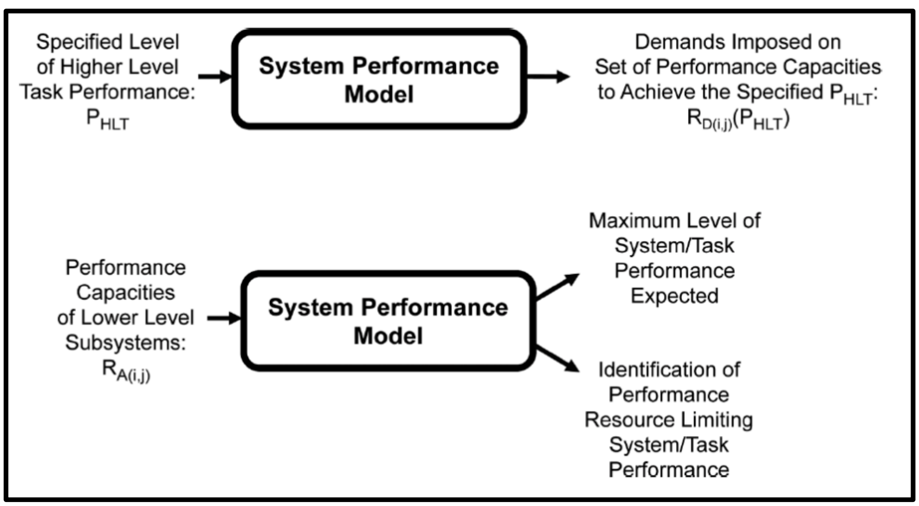


Figure 22. GSPT System Performance Model.
Source: Kondraske (2011, p. 237).

Economic theory grounds GSPT in terms of R_A and R_D. The relationship between the two establish the thresholds that define performance outcomes. Tvaryanas' (2010) claim that HSI resulted from the resource mismatches between human resources and system demands suggests it may be appropriate to proceed in terms of economics. GSPT defines HSI domain resources as BPRs and measures actual performance. This approach indicates a quantitative relationship can be established between the two variables. These characteristics enable system performance forecasting, a key HSI requirement. Additionally, the potential capacity to translate system data in terms of optimization provided reasonable incentive to evaluate this construct. GSPT may also improve HSI's communication with SE, in engineering terms, due to its system agnostic approach to input and output measurement. This initial review indicates GSPT/NCRA could satisfy the MBHSI-FRs. If so, improved HSI capacity to enable SE should be realized. Next, a description of GSPT/NCRA economic resource theory highlights specific elements critical to HSI resources.[11]

d. Economic Resource Theory Key Concepts

GSPT and MBHSI are founded on economic theory expressed as resource availability within subsystem *i* across DoP *j* to be greater than or equal to resource demand within subsystem *i* across DoP *j* as a condition of success. This relationship views availability and demand as things to be measured. This theory is expressed as:

$$R_{Ai,j} \geq R_{Di,j}$$

The unique system resource availability is an N-dimensional PCE expressed as:

$$\prod_{i=1,j=1}^{n,m} R_{A_{i,j}}$$

The R_A profile produced is a multiplicative (not additive) valuation approach that ensures all BPRs maintain their uniqueness and are not inappropriately traded, as they are truly different resources. In chemistry for example, each element is unique, one cannot trade hydrogen for oxygen. Resources of like kind are the only legal trades in terms of equal valuation. The additive models seen throughout the literature and in practice such as the ASVAB, the GRE, and others engage in additive models. For example, a GRE score

[11] For a comprehensive discussion of GSPT, see: Kondraske, (2011).

inclusive of a math score of 500 and a verbal score of 500 is considered a total score of 1,000. If another GRE score comprised of a math score of 300 and a verbal score of 700, it would still produce a total score of 1,000. Math and verbal resources are not the same. A multiplicative model would identify these differences as the scores would be 250,000 and 210,000, respectively.

R_D is a signal that is measured then assessed for position within or outside of the system PCE for performance insight expressed as:

$$\prod_{i=1,j=1}^{n,m} R_{D_{i,j}}$$

The expanded architecture expressed as:

$$(R_{A1,1} \geq R_{D1,1} \text{ AND } R_{A2,1} \geq R_{D2,1} \text{ AND } R_{A3,1} \geq R_{D3,1}\ldots)$$

If the HLT encapsulates the essential task elements, then the vector of BPR threshold values defines the system BPR resource profile.[12] Figure 23 illustrates this concept using a simple example involving only two BPRs. Resources are measured individually on each axis and the envelope is extrapolated. The system's PCE is then compared to the HLT BPR demand profile, also illustrated in Figure 23. If the HLT BPR demand profile is within the PCE, the system's performance R_A exceeds HLT R_D and at least minimally accepted performance should be expected. The two red dots illustrate the demand signals of two tasks. If the system's BPR demand profile is outside the PCE for some given level of performance, at least one of the system's performance resources will limit HLTp or result in failure. Figure 23 also demonstrates the malleability of one BPR by illustrating the potential training effect on the PCE. This is discussed in detail in Chapter VI.

[12] Note that unless all elements comprising the Table of Human Basic Elements of Performance are measured, the system resource profile is not defined completely and exhaustively. The Satisficing principle described in Simon's Bounded Rationality (Simon, 1955) suggests this is a rational approach. Additionally, the Pareto principle contends that measuring an important few BPRs that most significantly affect HLTp should suffice to enable satisficing performance forecast accuracy.

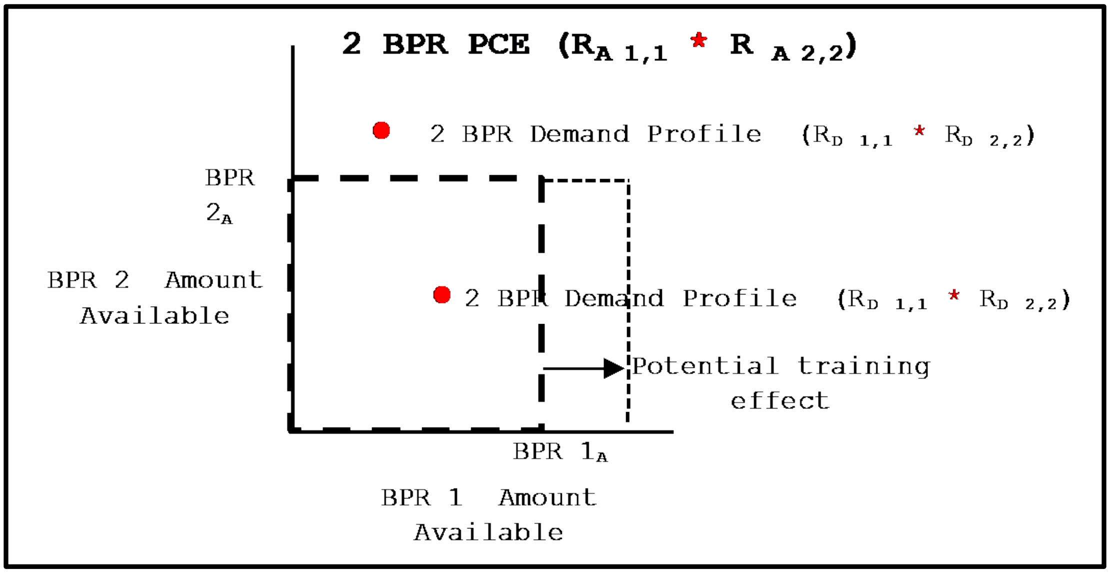


Figure 23. A Two-BPR PCE and Notional HLT Demand Profiles

e. Review of Relevant GSPT/NCRA Literature

A sample literature review of GSPT/NCRA contains performance forecasting studies executed across a spectrum of complex tasks including driving, surgical procedures, and leadership. NCRA outcome data demonstrate accurate performance forecasts across these studies where challenging performance measurement methods exist. GSPT/NCRA do not have a homogenous population requirement and seek to identify the relationship between a limiting resource and task performance.[13] Kondraske (2011) typically presents agreement between NCRA forecasts and actual HLTp using tabular views of accuracy allocated along bins of percentage.

f. Driving Studies

Two separate driving studies demonstrate excellent agreement between forecasted and observed performance. The first study titled "Prediction of Driving Performance Using Nonlinear Causal Resource Analysis" by Fischer et al. (2002) included 69 participants, ranging in age from 20 to 87 inclusive of various forms and stages of disease processes,

[13] For an extensive collection of GSPT/NCRA external references, see: *General Systems Performance Theory: Annotated Bibliography - External Sources.* A 2019 Technical Report highlighting publications citing GSPT and/or NCRA (George V. Kondraske, 2019).

demonstrated strong alignment with expert raters ($r = .83$) as shown in Figure 24. The study selected just 10 BPRs and was administered on a large hospital campus. The subjects were assessed using an analog scale ranging from 0–10. A similar driving study titled, "Sensory-Motor and Cognitive Tests Predict Driving Ability with Brain Disorders" by Innes et al. (2007) produced consistent results. They compared NCRA with other methods, suggesting NCRA was a promising approach. Specifically, "NCRA correctly classified 94% and 90% of referrals as on-road pass or fail" (Innes et al., 2007, p. 188).

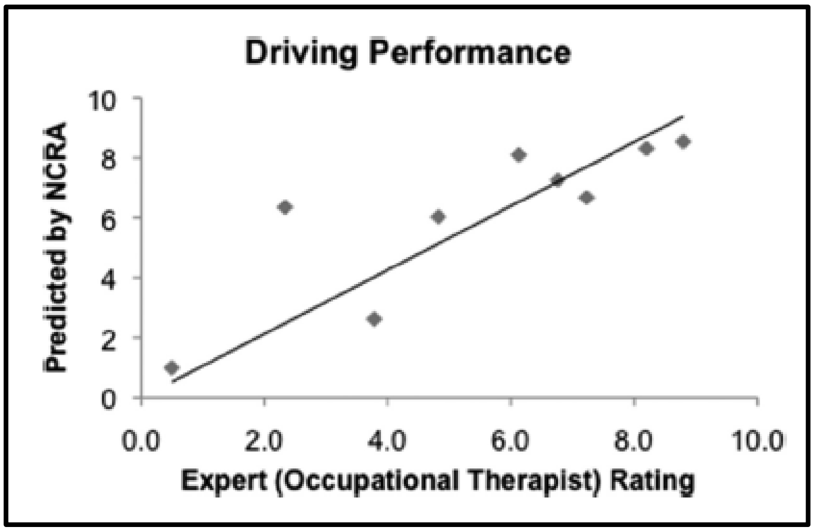


Figure 24. GSPT/NCRA Driving Study. Source: Fischer et al. (2002).

g. *A Surgical Study*

A surgical study also produced excellent evidence in support of NCRA's effectiveness in accurately predicting performance in a complex operational environment. The laparoscopic surgical study titled "Assessment of Basic Human Performance Resources Predicts Performance of Ureteroscopy" by Matsumoto et al. (2006) examined forecasted endoscopic surgical performance against expert raters. Potential in-residence selection, career counseling, and remediation therapy all exist in this domain to improve HLTp within the resource demands in the practice. Sixteen medical school students underwent virtual training and then executed the tasks on cadavers. The study measured 16 urology residents across 13 BPRs to build the forecast models. NCRA accurately predicted

10 of 16 medical students' performance within 15% of actual performance. Figure 25 illustrates the actual vs. forecasted performance. Mean ratings from three experienced surgeons on a 0–5 analog scale derived performance scores.

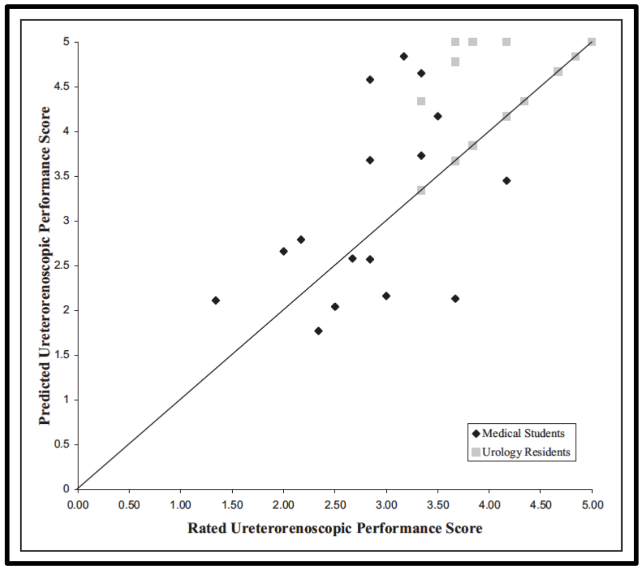


Figure 25. Actual vs. Predicted Ureterorenoscopic Performance Scores.
Source: Matsumoto et al. (2006).

h. A Leadership Study

Few operational tasks demonstrate the complexity of leadership. This particular study engaged GSPT/NCRA demonstrating 50.7% of cadet's leadership performance forecasts were within 20% of observed HLTp. In 2009, Bartone et al. published the study titled "Big Five Personality Factors, Hardiness, and Social Judgments as Predictors of Leader Performance." "The purpose of this paper [was] to evaluate the influence of psychological hardiness, social judgment, and 'Big Five' personality dimensions on leader performance in U.S. military academy cadets at West Point" (Bartone et al., 2009, p. 1). The study evaluated cadets during academic semesters and field training exercises while they were at West Point. Supervisor-generated leadership grades during their tenure served as HLTp data. Four hundred seventy-five cadet data sets were included in the National Defense University (NDU) study. Table 1 provides detailed forecast accuracy data.

Table 1. National Defense University Leadership Performance Forecasts Using NCRA. Source: Bartone et al. (2009).

Within Range Specified			Cumulative		
% Difference (Predicted vs. Measured) Range	No. Subjects	Percentage of Subjects	% Difference (Predicted vs. Measured) Range	No. Subjects	Percentage of Subjects
<5%	73	15.4	<5%	73	15.4
≥5 & <10%	47	9.9	<10%	120	25.3
≥10 & <15%	66	13.9	<15%	186	39.2
≥15 & <20%	55	11.6	<20%	241	50.7
≥20 & <25%	42	8.8	<25%	283	59.6
≥25 & <30%	46	9.7	<30%	329	69.3
≥30 & <35%	34	7.2	<35%	363	76.4
≥35 & <40%	31	6.5	<40%	394	82.9
≥40%	81	17.1	≥40%	475	100.0

In summary, the data from all four studies demonstrate impressive forecast agreement with actual HLTp, especially given the contexts of the studies. NCRA accuracy improved as the number of BPRs increase. Typically, NCRA over predicted performance, suggesting that additional limiting BPRs are present in the task environment. These data are feedback legitimizing continued search for appropriate BPRs. Furthermore, NCRA data support decisions regarding individuals, not just populations. This level of precision in the human performance arena is elusive. The success demonstrated by NCRA in terms of establishing a cause-and-effect relationship between BPRs and HLTp suggests a potential to improve the capacity of HSI may be possible.

3. Summary

This chapter has shown that DOD HSI policy DoDD 5000.01, DoDI 5000.02 (to include the draft DOD 5000.PR "Human Systems Integration in Defense Acquisition," the Defense Acquisition Guide, as well as the requirements established by HSI containing systems, all converge on the purpose of MBHSI as defined by this dissertation's set of MBHSI-FRs. GSPT/NCRA, previously identified by the DSOC HSI Task Force as a potentially beneficial theoretic framework, is able to model the human system in terms of resources and performance using actual measurement, thereby satisfying the MBHSI-FRs. Chapter III establishes MBHSI using the reference of MBSE, re-characterizes the seven HSI domains in terms of GSPT, and introduces the dissertation study.

THIS PAGE INTENTIONALLY LEFT BLANK

III. MODEL-BASED HUMAN SYSTEMS INTEGRATION

> Dreamers envision our future: pragmatists build it. Solving the right problem in a right way moves mankind forward. Cleverness and ingeniousness aside, solving the right problem in a wrong way or the wrong problem in a right way demonstrates a crucial systems problem—achieving an integrative end state.
>
> —Gary O. Langford, *Engineering Systems Integration*

A. PURPOSE

This chapter serves two purposes:

1. To define Model-Based Human Systems Integration (MBHSI).

2. To describe the dissertation laboratory study.

Figure 26 illustrates the Chapter III roadmap. Each node on the roadmap is accomplished sequentially with the exception of "achieve criteria for success." Chapters IV-VI (Projects I-III) accomplish this final node.

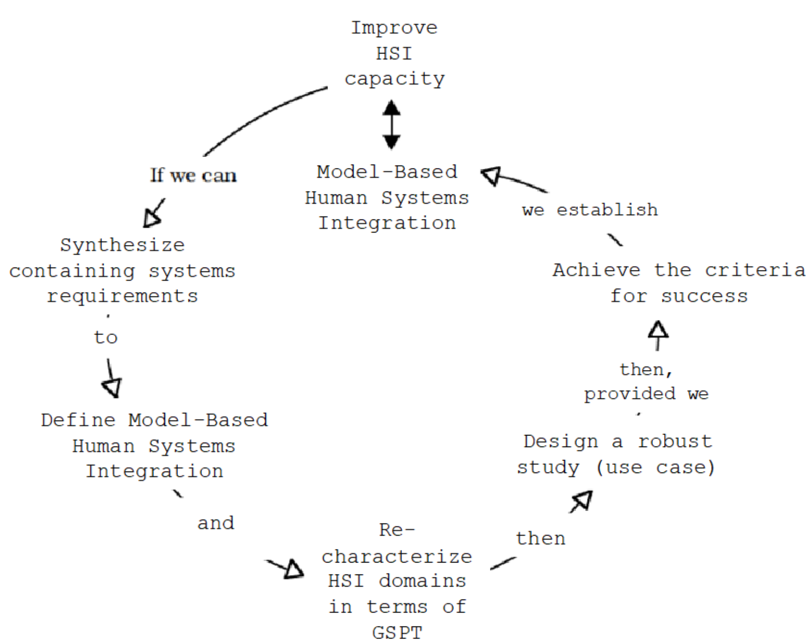


Figure 26. Chapter III Roadmap. Adapted from Hitchins (1992).

B. BACKGROUND

The purpose of this new concept, MBHSI, is to improve HSI's capacity to enable SE. The goal of MBHSI is to align with the MBHSI-FRs using MBSE as a reference point. MBHSI objectively defines and leverages the quantitative relationships between HSI domain resources and system-level performance. In Chapters I and II, the researcher raised the challenge of identifying a theoretical perspective for defining HSI variables to include determining their values. This necessary precondition for modeling links these variables to the outcome of TSP. The 2010 DSOC HSI Task Force report offered GSPT/NCRA as a possible unifying HSI theoretical approach that has not yet been investigated. This theoretical perspective can be nested within the containing system of MBHSI as an enabling system because it appears to support the required HSI outputs and this dissertation's MBHSI-FRs.

C. DEFINITION

MBHSI emerged by seeking to "diminish uncertainty and related disorder [that can be] generated by an inward-oriented system [and complexity] by creating a higher and broader more general concept" (Boyd, 1976, p. 7). The goal was to build a new concept, using MBSE as a reference point, to improve HSI's capacity to deliver requisite outputs to SE. The researcher's definition of MBHSI resulted from the synthesis of the DOD HSI objectives, policies, and guidance; definitions offered by INCOSE and Tvaryanas (2010); HSI systemic diagnoses proffered in the past; and HSI's impetus—complexity:

> MBHSI is an essential, model-based, and integrative process that reliably addresses complexity in terms of resource economics while enabling the SE practice. It applies GSPT and NCRA to model and forecast the quantitative relationships between HSI domain resources and system-level performance, targeting the chronic HSI trade space problem and the original objective of HSI, optimization. Finally, it seeks to communicate its engineering and program management value in engineering terms.

The containing systems requirements, identified in Chapter II, inform the purpose of MBHSI as an enabling system to SE. In Chapter II, the concept of systems hierarchy and required outputs partially explained the requirements relationships between systems.

Figure 27 illustrates these defining relationships within the MBHSI system hierarchy; specifically, the requirements of each containing system reflect enabling system outputs.

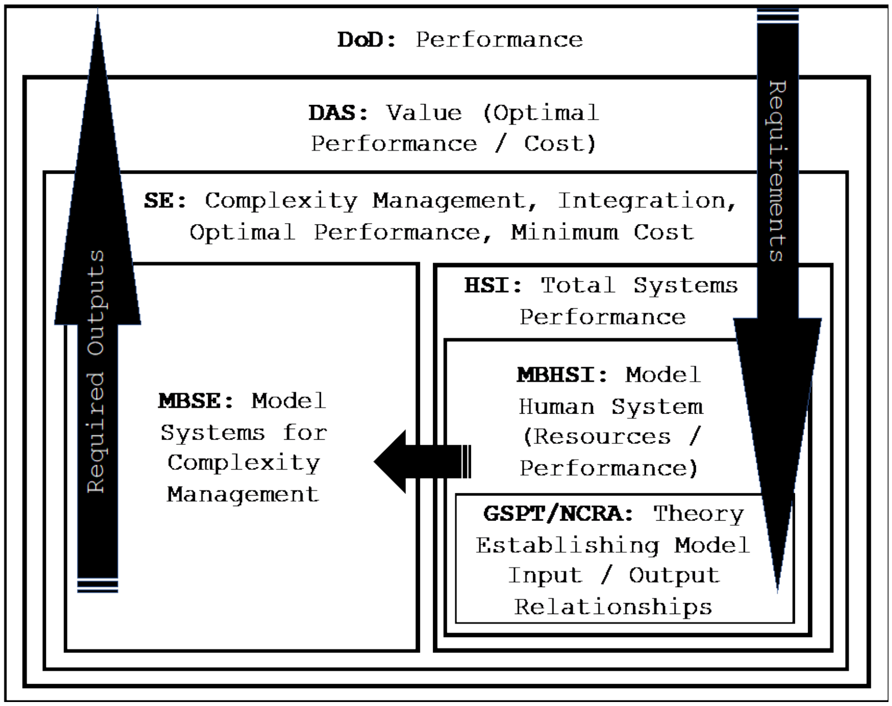


Figure 27. MBHSI System Hierarchy Mapped to Requirements. Adapted from Hitchens (1992, p. 10).

The foundational nature of a unifying theoretical construct and its potential influence on containing systems is an important concept. The relationship between requirements and outputs is limited by the capacity of the enabling system to adhere to the specified requirements. For example, MBSE's outputs are limited by its enabling systems. Should HSI fail to appropriately model the human system in useful terms, the human will continue to remain an afterthought in systems design (Madni & Orellana, 2018). If a system fails to meet requirements, then it cannot produce the appropriate outputs. The added emphasis in INCOSE's Vision 2025 regarding theoretical foundations emphasizes this critical point (INCOSE, 2014). Therefore, in order to facilitate required outputs to MBHSI, model inputs must be characterized in terms of GSPT.

D. HSI DOMAINS IN TERMS OF GSPT

Early exploration of GSPT as a theoretical foundation for MBHSI illuminated the need to re-characterize the HSI domain architecture in terms of GSPT. The purpose of this critical conversion is to provide a constructive bridge from HSI to MBHSI while maintaining the original intent of HSI and adhering to DOD HSI policy (requirements). DOD HSI policy acknowledges seven domains of HSI; therefore, MBHSI must achieve this same requirement. The practitioner will notice the contrast between current domain definitions and how MBHSI engages them. How MBHSI conceptualizes and measures the HSI domains are key distinctions. MBHSI engages the resource economic model described in Chapter II and prioritizes measurement of resource available (R_A), resource demand (R_D), and high-level task performance (HLTp).[14] The domains must be interpreted within the GSPT context.

A few simple examples illustrate this important distinction:

- A manpower increase results in an increase in system-level PCE, which means improved system R_A.
- In the case where inadequate system Human Factors Engineering (HFE) drives down performance, a higher system R_D results in fewer personnel PCEs encapsulating the demand signal of the task.
- In the event that HLTp thresholds are increased, improved HFE and/or personnel with higher R_A will be necessary.
- In a situation where the R_D is high enough, short- and long-term PCE implications are expected (i.e., acute medical problems and/or chronic Veterans Affairs [VA] disabilities).

[14] Recall that nearly all humans have the same BPRs; it is the differences in capacity thresholds that distinguish individuals.

- Training or occupational health, if applied correctly, can expand or rehabilitate an individual PCE supporting non-medical and medical readiness.

MBHSI views HSI domains through the lenses of R_A, R_D, and HLTp. These lenses bring focus to measurement of, and understanding about, the relationships at play between HSI domain resources and system-level performance. This understanding facilitates defending, preserving, and expanding the human-system PCE and smartly reducing R_D where feasible.

Every task demands resources. When R_D exceeds R_A, near misses, accidents, failure, drift from stability, compromised resilience, injury, permanent damage, or even death can occur. Some of the HSI domains measure both R_A and R_D of the human system to sustain, optimize, and enhance the individual and collective PCE. These re-characterizations of HSI domains communicate not only internally within the practice but also externally with MBHSI containing systems. The re-characterizations also might communicate externally to the warfighter, training systems, personnel systems, occupational medicine, safety centers, human performance, academia, epidemiology, and the VA.

Following are DOD-recognized HSI domains after a re-characterization in terms of GSPT:

Resource Availability—R_A

- **Personnel: Individual human resources (R_A) to be defended, preserved, and expanded**
    - The starting point in terms of both malleable and non-malleable basic performance resources (BPR)
    - BPR identification in user population PCEs
    - The individual PCE for each person

Resource Availability—R_A

- **Manpower: System R_A**
  - Volume of PCEs in DOD specialties, organizations, or teams
  - Volume of system PCE

Potential Resource Availability—R_A

- **Training: PCE Expansion**
  - Set of malleable BPRs in user-population PCEs
  - PCE maintenance
  - Informed by MBHSI training effectiveness analytics (R_D)

Resource Efficiencies or Drain on PCE (Design Dependent)—R_D or R_A

- **Human Factors Engineering: R_D Expansion or Compression**
  - Defines levels/types of required BPRs given an HLTp
  - Defines task R_D (resource demand function [RDF] shape)
  - Adjusts task analyses
  - Proceeds only post-R_D and R_A diagnostic evaluation
  - Targets treatment of a system to relieve limiting BPRs (evidence-based)

Acute Resource Demands or Preservation—R_D or R_A

- **Force Protection/Survivability: Minimize R_A Loss**
  - Margins from edges of PCEs and task (defined by limiting BPRs and critical phases of a task)
  - Survivability margins required to ensure maximum margin is present across a pre-determined set of BPRs

Acute, Chronic, and Longitudinal PCE Compression or Regeneration—R_D or R_A

- **Habitability: Smart PCE Regeneration**
  - R_D signal or R_A injection from the living environment
  - Regenerates, maintains, or compresses PCE volume

- Regeneration or compression of PCE is considered non-static; it evolves or progresses over time.
- These demands and/or regeneration can be acute or chronic, and/or longitudinal in nature.

Habitability R_D should be minimized in an effort to maintain the system PCE. To the extent they do not, resources (BPRs) are wasted. Routine measurement of the PCE and HLTp defines the R_D of the HLT over time. R_D of the human system to ensure adequate R_A to meet the demand should also be measured (i.e., required quantity and quality of sleep).

Acute, Chronic, and Longitudinal Demands, Prevention, Diagnoses, and/or Rehabilitation—R_D or R_A

- **Occupational Health and Safety: Resilience Engineering**
  - Environmental and task R_D surrounding the HLT
  - Safety is some correlation of the R_D / R_A relationship
  - Includes preparation for the HLT, the HLT, and the recovery from the HLT
  - R_D are acute, chronic, and/or longitudinal in nature.
  - Initial, scheduled, and as-needed assessments serve as quantitative readiness measures.
  - R_D of the human system to ensure adequate R_A to meet the demand should be measured (i.e., oxygen, hydration, nutrition, sleep)
  - Drift from dynamic stability within a system can be detected here. The provider can then quantify when a human system has the capacity to return to duty.

The occupational health portion should be transparent to the VA. Also, engagement with epidemiologists as feedback and evidence mechanisms to the parent organization may add value. VA data could be used to inform the boundaries of the HLT R_D signal.

In terms of safety, an accident investigation might conduct a root cause analysis (RCA) using PCE and HLT demand signals at the time of the event. Demands that exceed availability or where resource availability was expected but not present could be causal in the event. Both malleable BPRs and dimensions of performance (DoP) demands of the system inform recommended preventative and corrective measures.

In summary, this re-characterization of HSI domains in terms of GSPT supports a new concept of HSI while preserving the historical intent of the practice and adhering to policy requirements. MBHSI is serious about measurement, transparency, evidence-based methods, and external communication with appropriate systems. MBHSI seeks to interface with SE by establishing the human system in terms of R_A data and R_D impact on system performance (HLTp). This new concept suggests that an evidence-based approach to HSI may improve the capacity for HSI to enable SE.

E. THE RESEARCH STUDY

To demonstrate application of GSPT/NCRA as a potential unifying construct for MBHSI, the researcher completed the laboratory research study titled "Model-Based Human Systems Integration Using General Systems Performance Theory." The study included two major measurement components:

- R_A: Nineteen cognitive and psychomotor BPR capacity measurements (inputs)
- R_D: HLTp measurement using a simulated general aviation Instrument Landing System (ILS) approach (outputs)

The study consisted of three iteratively built research projects:

- Project I (Chapter IV): MBHSI Performance Forecast Model Development
- Project II (Chapter V): MBHSI Resource Demand Functions and System Performance Forecasts
- Project III (Chapter VI): MBHSI Trade Space

Figure 28 maps the projects and the MBHSI-FRs (criteria for success). In this case, the MBHSI-FRs deductively flow down to the core requirement of defining HSI domain resources as model inputs. During execution, each build on the prior as outputs emerge, culminating in the identification of optimal system solutions.

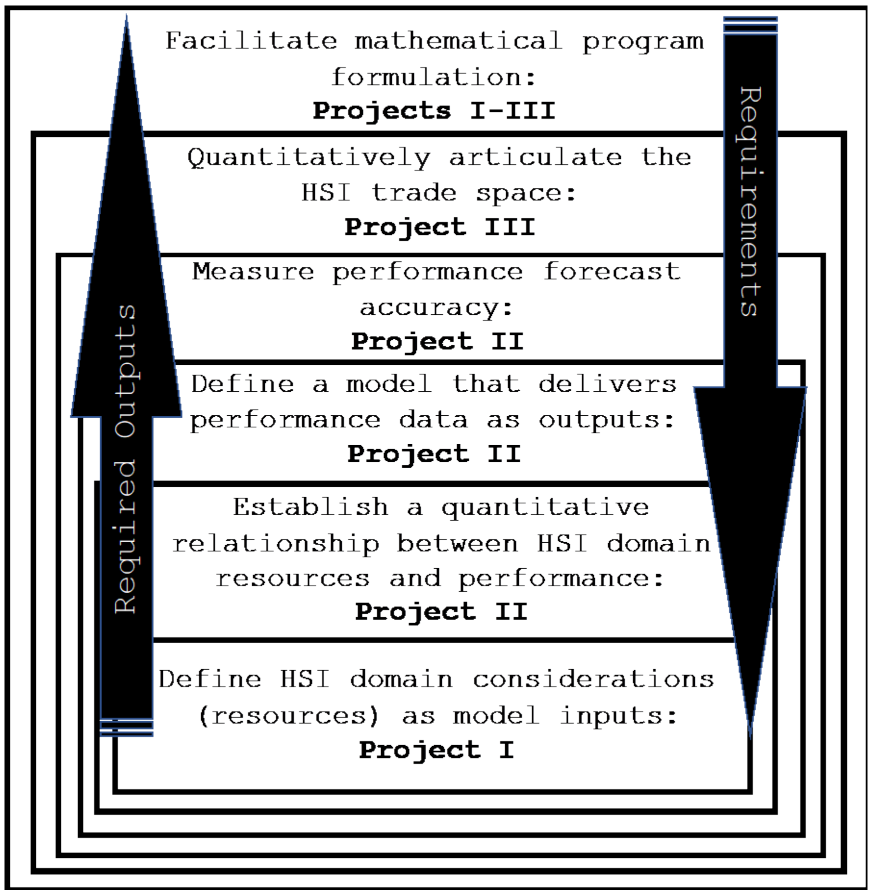


Figure 28. Research Projects Mapped to HSI Functional Requirements

The single study served as a use case to the three iteratively built research projects. This portion of Chapter III provides the study introduction and methods pertaining to all three projects. Information specific to a particular project can be found within that project. This approach modularizes the dissertation, reduces repetition, and creates a reference point for the reader regarding the study. Figure 29 illustrates this approach including the criteria for success as a requirement for MBHSI.

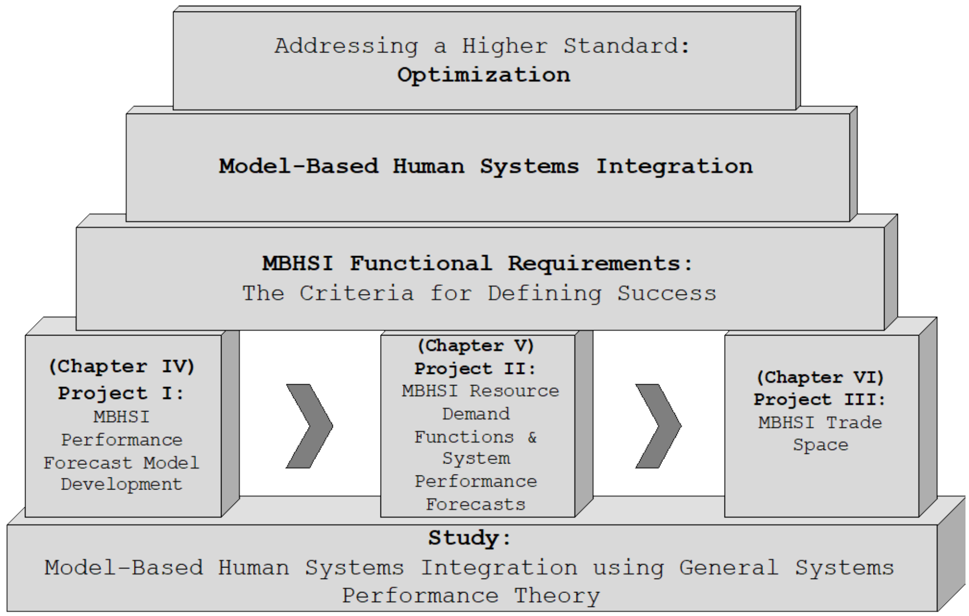


Figure 29. MBHSI Study Approach

1. Problem Statement

DOD HSI lacks a theoretical perspective that bridges domain resources with system performance. Therefore, as a contributing discipline to SE/MBSE, HSI lacks a complementary model-based approach to effectively integrate the human system into MBSE system architectures. In order to enable MBSE, HSI must develop complementary methods for modeling, forecasting, executing trades, and optimization, according to DoDI 5000.02PR (draft).

2. Objectives

As stated in Chapter I, the objective of this dissertation is to empirically define the relationship between human performance resources and system-level performance for the purpose of increasing the capacity of HSI to enable SE. This research study examined the efficacy of GSPT/NCRA to:

- model the human system in terms of performance resource capacity (BPRs) and system-level performance (HLTp)
- execute accurate performance forecasts

- explore resource expansion and reduction influence on HLTp
- inform quantitative trades across domains
- support formulation of mathematical programs leading to optimal solutions

3. Hypothesis

The overarching hypothesis adopted for the present study are:

- **H_0:** Current (state-of-the-art) HSI does not adequately deliver the required outputs to its containing systems (Condition A: Atheoretical).
- **H_A:** HSI can be reframed in terms of GSPT and NCRA to improve adherence with containing system requirements (Condition B: Theoretical).

4. Methodology

"Model-Based Human Systems Integration using General Systems Performance Theory" was approved by the NPS Institutional Review Board (IRB) and met exemption category 3 in accordance with 32 CFR 219.101(b). The NPS IRB office confirmed completion of CITI Research Ethics Training. Though the IRB did not direct informed consent, consent was obtained for each participant as a best practice. Appendix A contains a copy of the study consent form. All data were stored in accordance with the approved IRB protocol.

5. Subjects

Sixty-six individuals responded to recruitment primarily from NPS and the Center for Homeland Defense and Security (CHDS) located at NPS. The study retained 97 percent of the data sets (64/66).[15] The IRB-approved protocol authorized up to nine non-federal

[15] Only two participant data sets were excluded: one for bizarre HLTp data that suggested a strong preoccupation with one performance metric (Airspeed) at the expense of the other metrics and one that demonstrated a language barrier that suggested confusion with instructions, assessment, and task objectives.

government employees; all nine were used. Participant recruitment included flyers (see Appendix B), personal interaction, and student solicitation with instructor approval. Female participants accounted for 36 percent of the participants (23/64). Ages ranged from 21 to 63 with mean *and* median ages of 41.5 years. Mean gaming experience averaged less than one hour per month. All but two participants held at least a bachelor's degree. Nineteen held master's degrees. Eleven reported doctoral degrees. Participant occupational backgrounds demonstrated excellent heterogeneity. They included, but were not limited to, tenured professors, firefighters, naval surface warfare officers, religious lobbyists, secret service officers, physicians, border patrol agents, stay-at-home parents, a retired admiral, civilian college students, FBI agents, USMC infantrymen, university security officials, FEMA personnel, registered nurses, major seaport entry officials, statisticians, and engineers. Study length ranged from 1.5 to 2.0 hours, depending on participant track assignment. Eligibility criteria included subjects must be over 18 years old, a federal government employee (except for nine participants), and have zero time piloting any aircraft. Eight separate pilot-testers assisted with BPR collection protocols and setting the HLT demand signal appropriate for a novice pilot. These pilot-testers included novice and experienced aviators.

This study warrants a note regarding statistical power analysis. The GSPT/NCRA construct is threshold based. Because GSPT/NCRA uses the RDF approach to forecast performance, the commonly used sample size formulas (power analysis) are not appropriate. The GSPT/NCRA approach to performance analytics requires a sample size sufficient to define the RDF for HLTp, that reliably generates accurate forecasts. NCRA forecast accuracy prefers a subject population with a uniform R_A distribution across an appropriate HLTp range. This will likely result in a higher quality (more accurate) RDF (Kondraske, 1998). Therefore, GSPT/NCRA prefers a heterogeneous population sample.

6. Research Design and Variables

Table 2 summarizes the research design, variables, and overarching conditional expectations.

Table 2. Research Design and Variables

Type of study: Model-Building
• Human subjects research • Forecast modeling via simulated task in a virtual environment
Design: Mixed
• Pretest measurement • Pre-posttest measurement
Independent Variables:
• Training and/or Automation
Dependent Variables:
• Test scores and HLTp
Conditional Expectations:
• $E[HLT_P \mid R_A]$ • $E[R_A \mid HLT_P]$

7. Instruments

Demographic data collected from participants followed the signing of informed consent for the study. Data included: age, gender, education level, gaming experience, flight experience, and physical anomalies (vision and hearing). Appendix C contains a copy of the demographic data collection form. The BPR measurement apparatus included a cognitive and psychomotor test battery of eight different tests, generating nineteen BPR measurements. Project I (Chapter IV) details the instruments used to collect BPR (input) data. Table 3 documents the flight simulation equipment used to collect HLTp data. Figure 30 illustrates the HLT.

Table 3. HLT Equipment Inventory

HLT Equipment
Software
X-Plane 11
Hardware
Hewlett Packard Z4 desktop computer
LG 60-inch LCD flat screen
Flight Controls
Logitech flight simulator yoke system
Logitech flight simulator rudder pedals

Figure 30. MBHSI HLT Instruments

8. High-Level Task Design

Pilot testing helped develop a specific flight configuration ensuring each participant received a standardized HLT. The final HLT configuration was a simulated extended final ILS approach to Runway (RWY) 34R at Seattle-Tacoma International Airport (KSEA) in a Cessna-172, starting at just over nine miles from the runway. Visual meteorological conditions (daylight, some high clouds, and no wind) describe the weather conditions. No other air traffic was present in the profile. The boundaries of the scored portion of the approach were set six miles out to one mile from the runway, creating a five-mile ILS approach. The first three miles of each HLT served as a participant orientation and coaching by the researcher, a licensed pilot. Figure 31 provides a portion of the KSEA approach plate illustrating the portion of the approach scored for the HLT. Specifically, the scope of the HLT was the shaded region extending out from the threshold of the runway from one to six miles. Appendix D provides the entire current KSEA RWY 34R approach plate.

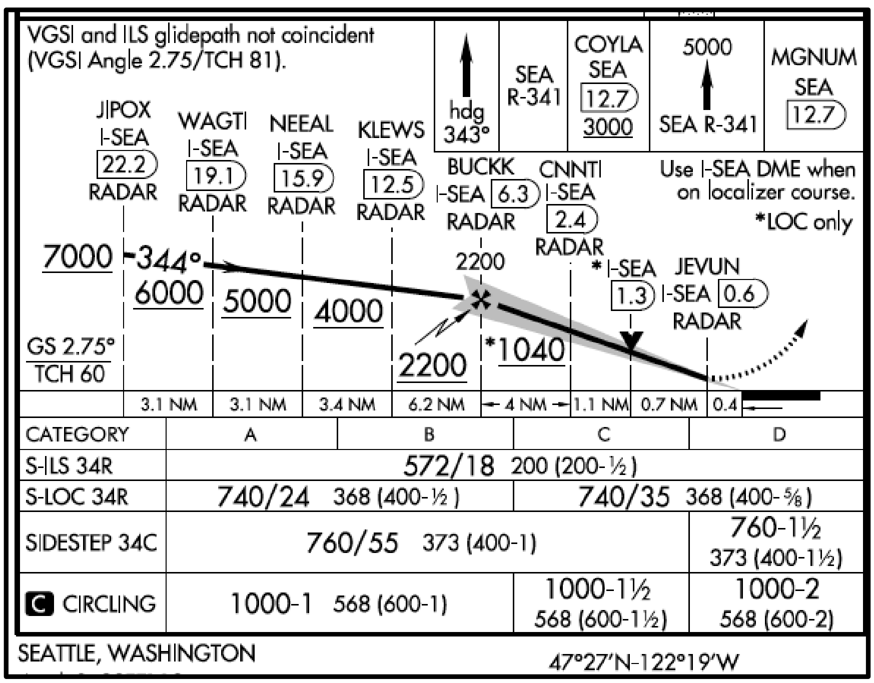


Figure 31. KSEA RWY 34R Approach Plate Segment

Performance data collection occurred at 1Hz intervals (every second) and included Airspeed and Course Deviation Indications (CDI) for both Course and Glideslope. Data collection began after the three-mile period following unfreezing the simulation. Data collection ended at Decision Height (DH), one mile from the threshold of the runway. The landing was not scored. Each participant received one practice approach (HLT 0, not assessed) and two scored approaches (HLT 1, assessed) and (HLT 2, assessed). The HLT was decomposed into three lower-level sub-tasks. This demonstrates MBHSI analytical capacity at different levels of the HLT hierarchy. These sub-HLTs coincide with the performance metrics of the HLT. The three sub-HLTs are

1. Airspeed
2. Course (CDI-H)
3. Glideslope (CDI-V)

Variations of HLT 2 existed for some of the participants. Project III in Chapter VI documents the specific methods and results pertaining to this portion of the study. An overview of the HLT configurations and various participant tracks provide important information regarding the study. Ten participants received flight and instrument training via Jeppesen professionally developed videos between their HLT 1 and HLT 2 approaches (Track B). This served to change the shape of the participant PCE while leaving the HLT static. Repeated measurement of HLTp and posttest knowledge BPR measurement quantified training effectiveness. Seven participants had the Course portion of the task automated (Track C.1), along with another seven that had the Glideslope portion of the task automated (Track C.2) between their HLT 1 and HLT 2 approaches. These 14 participants accomplished the HLT with a reduced demand signal. Repeated measurement of HLTp following a reduction of the HLT R_D quantified effectiveness of demand reduction across system configuration types. Participant tracks A, B, and C were defined as:

- Track A—Control
    - N = 40[16]
    - Completed all BPR data collection to include knowledge pre-test for baseline data
    - Completed HLT 0 (practice flight)
    - Completed HLT 1 (scored HLT)
    - Received no training
    - Completed HLT 2 (second scored HLT) to establish learning effect data

[16] All 64 participants HLT 1 data contributed to Projects I and II.

- Track B—Training
    - N = 10
    - Completed all BPR data collection to include knowledge pre-test for baseline data
    - Completed HLT 0 (practice flight)
    - Completed HLT 1 (scored HLT)
    - Received professional training
    - Completed HLT 2 (second scored HLT) to establish training effectiveness data (HLTp), a modified HLTp score
    - Completed knowledge posttest to establish PCE expansion data (BPR), a modified PCE

- Track C.1—Course Automated
    - N = 7
    - Completed all BPR data collection to include knowledge pre-test for baseline data
    - Completed HLT 0 (practice flight)
    - Completed HLT 1 (scored HLT)
    - Received no training
    - Completed HLT 2 (second scored HLT) with Course (CDI-H) automated via autopilot to observe a modified HLTp score

- Track C.2—Glideslope Automated
    - N = 7
    - Completed all BPR data collection to include knowledge pre-test for baseline data
    - Completed HLT 0 (practice flight)
    - Completed HLT 1 (scored HLT)
    - Received no training
    - Completed HLT 2 (second scored HLT) with Glideslope (CDI-V) automated via autopilot to observe a modified HLTp score

9. Procedures

One participant at a time completed the study in the MOVES Institute Laboratory at NPS. Participants reviewed and signed consent forms, received study information, completed demographic data sheets, then provided BPR capacity measurements and executed the HLT as described. Specific study procedures are detailed in Projects I-III. Finally, each participant received a certificate of appreciation showing their PCE and HLTp scores. Appendices E and F contain copies of the procedural checklists, protocol, and certificate of appreciation.

F. SUMMARY

This chapter finalized a form of creative induction defining a new concept of HSI, MBHSI (Boyd, 1976). The definition of this new concept synthesizes HSI objectives, requirements, policies, guidance, definitions, systemic diagnoses, and complexity while using MBSE as a reference point. Policy requires MBHSI adherence to DOD-recognized HSI domains, yet GSPT requires appropriate model inputs. Therefore, a necessary re-characterization of HSI domains satisfies these dual requirements. This newly minted concept is now primed to deliver required outputs to MBSE as defined by this dissertation. This dissertation's laboratory study demonstrates application of MBHSI and the stated MBHSI-FRs define the criteria for success. Chapters IV-VI (Projects I-III) serve to document evidence that evaluate the veracity of MBHSI's potential to increase the capacity of HSI:

- Project I defines model inputs and outputs by measurement. The relationship between HSI inputs and outputs emerge quantitatively in the GSPT models.

- Project II establishes the RDF and explores the utility of the novel MBHSI code logic. This section includes evidence of RDF development and encouraging forecast accuracies.

- Project III explores PCE expansion and R_D reduction (improved HFE) and the impacts on HLTp. This demonstrates constraint relaxation or restriction in defining feasible solution sets. The pre-posttest design demonstrates quantitative insights into training effectiveness and system configurations.

A comprehensive discussion follows Chapter VII, including an appraisal of MBHSI, documenting evidence of success per the defined criteria: the MBHSI-FRs.

IV. RESEARCH PROJECT I: MBHSI PERFORMANCE FORECAST MODEL DEVELOPMENT

Only those who take leisurely what the people of the world are busy about can be busy about what the people of the world take leisurely.

—Taoist Maxim

A. INTRODUCTION

The purpose of Project I was to test GSPT efficacy to establish relationships between HSI resources (inputs) and performance (outputs). The criterion for success was the development of scatterplot models that reflect GSPT characteristics. If these relationships are successfully established, then the output requirement of GSPT as a contained system within MBHSI is achieved, as shown in Figure 32. In Project II, these relationships define the nonlinear threshold-based envelopes conceptualized by NCRA, the RDFs. NCRA uses these RDFs to forecast system performance.[17]

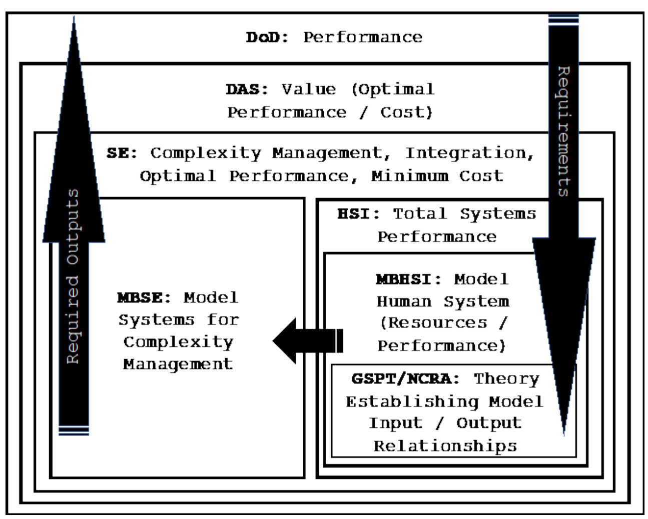


Figure 32. MBHSI Requirements Hierarchy and Related Outputs. Adapted from Hitchens (1992, p. 10).

[17] For a comprehensive discussion of GSPT and NCRA, see: Kondraske (2011).

The second purpose was to satisfy the core MBHSI-FR, *define HSI domain considerations (resources) as inputs*, which is illustrated in Figure 33.

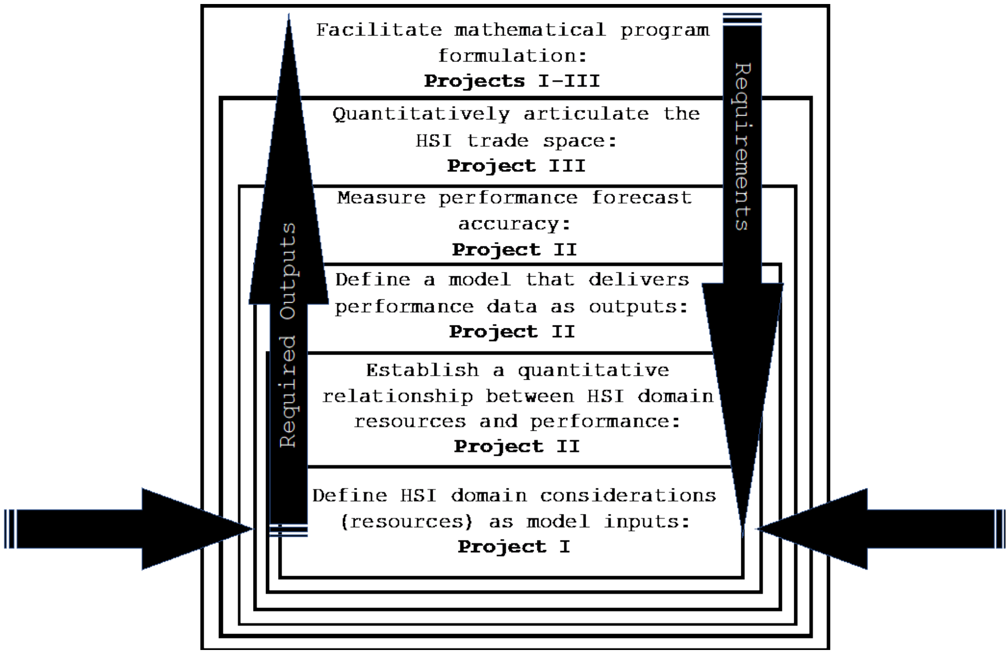


Figure 33. Project I Mapped to MBHSI-FR

The goal was to demonstrate accurate R_A and HLTp measurement. Prior to measurement, a definition of the HLT and initial satisfactory performance is required. This requirement is not unique to GSPT or MBHSI; it is a critical tenet of systems thinking. These measures model the relationship between R_A and HLTp, operationalizing MBHSI in terms of GSPT. The establishment of these relationships develop the MBHSI performance forecast models. These models serve as the critical prerequisite for the establishment of the RDFs in terms of NCRA. The RDFs developed in Project II are the backbone of MBHSI because they quantify resource demand (R_D) for a given task at various levels of HLTp. Figure 34 illustrates the roadmap for Project I.

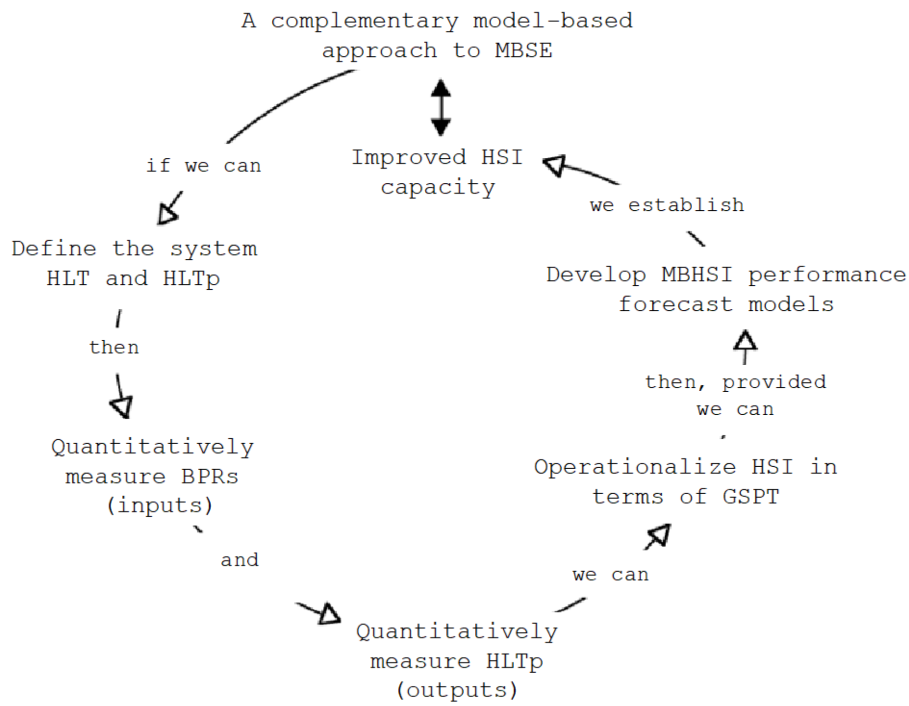


Figure 34. Project I Roadmap. Adapted from Hitchins (1992).

1. Problem Statement

As a contributing discipline to SE, HSI lacks an accurate and reliable methodology for modeling the human performance resource and system-level performance relationship. This research project demonstrated a theoretically-based process to construct accurate system performance forecast models. This methodology contributes to HSI containing systems and establishes a complementary HSI approach to MBSE. In this dissertation, these models are shown to address the HSI performance forecasting, trade space, and optimization challenges that persist in the DOD.

2. Project I Objective

The objective of Project I was to quantitatively measure GSPT model inputs (BPRs) and outputs (HLTp) for all participants. MBHSI models built in this project provide foundational data supporting the other three research projects, which are discussed later:

3. Hypothesis

GSPT can reliably model the human system to establish relationships between system inputs (BPRs) and outputs (HLTp).

4. Testing the Hypothesis

Success is defined as a set of GSPT representative scatterplots sufficient to engage NCRA.

B. OVERVIEW

Chapter III documented the study sample, overarching research design, variables, and HLT instruments to modularize this dissertation's three iteratively built projects. This chapter documents Project I's specific methodology, research design, variables, BPR instruments, procedures, and analytics.

C. METHODOLOGY

The model-building study measured resource capacities across a set of 19 cognitive and psychomotor BPRs thought to limit HLTp, as described in Chapter III. Each BPR measurement sought maximum performance by participants to build their 19-dimension PCEs in accordance with GSPT practices. The highest scores for each BPR measurement defined capacity for each participant in the study. The study also measured novice pilot HLTp (outputs) using X-Plane 11 to simulate an ILS approach in a single-engine fixed-wing Cessna 172 under visual flight rules (VFR) conditions. Measurement occurred from six miles out all the way to decision height (DH), a five-mile precision approach. Performance measurements spanned three sub-HLTs (Airspeed, Course, and Glideslope). A relaxed Federal Aviation Administration (FAA)-referenced scoring rubric quantified HLTp using X-Plane 11 data output functions. Each participant received one practice approach (HLT 0) with instruction, then completed a second approach (HLT 1) for score. Landings were attempted, but not scored. Project I models include those with and without performance penalties enforced. The input and output data constructed the GSPT scatterplot models using R. Project I presents a total of 136 different models including HLT 1 and the three sub-HLTs, with and without penalties enforced across 17 of the 19 BPRs.

The two knowledge BPRs did not demonstrate appropriate GSPT characteristics as expected. Novice participants (assumed zero BPR) taking a primarily multiple-choice test (demonstrated capacity due to chance) yielded models consistent with normally distributed scatterplots. This predicted issue is discussed in detail in the results section of this project and in Project III. The GSPT models complete the prerequisite requirements for Project II: MBHSI Resource Demand Functions and System Performance Forecasts.

D. RESEARCH DESIGN AND VARIABLES

Project I was a within subjects, pretest measurement, model-building laboratory study using human subjects. The study inputs/outputs included HLTp and BPRs, respectively.

E. INSTRUMENTS

The two major measurements in this study, BPRs (inputs) and HLTp (outputs) required a comprehensive collection of instruments and measurement protocols. Chapter III documented the HLT instruments. This section focuses on BPR data collection. The first step included a review of the task guided by O*NET (O*NET, n.d.) and Fleishman's *Handbook of Human Abilities* (Fleishman, 1992). Pilot testing prior to executing the study fine-tuned the specific instrument protocol.

An O*NET report for commercial pilots produced a prioritized list of tasks, skills, and abilities. Only highly rated items (those thought to be of a limiting nature) informed BPR selection. Table 4 summarizes the top-ranked O*NET results.

Table 4. Highly Ranked O*NET Commercial Pilot Profile

O*NET (Pilot)	Ranking	Definition
Task	95/100	Use of instrumentation to pilot an aircraft when visibility is poor
Skill	91/100	Controlling operations of equipment or systems
Abilities		
Control Precision	81/100	Quick and repeated adjustment to controls of a machine to exact positions
Problem Sensitivity	78/100	Problem recognition

Fleishman (1992) assisted in the identification of the initial list of BPRs for consideration. As expected, there was strong agreement between O*NET and *Fleishman's Handbook of Human Abilities* for piloting an aircraft due to O*NET's reference of Fleishman (1992). Table 5 provides a prioritized list of BPR measurement instruments. Two of the BPR instruments generate multiple separate BPR measurements. The multi-limb coordination test generated speed, accuracy, and throughput BPRs. The Automated Neuropsychological Assessment Metrics (ANAM) switching task generated speed, accuracy, and throughput BPR data for math, spatial orientation, and switching tests. In total, the nine instruments listed produced 19 measurable BPRs. Seventeen demonstrated strong GSPT characteristics as described in Chapter III, and were used in Project II.

Table 5. MBHSI Project I BPR Instrument Inventory

BPR Instruments
1. Rotary Pursuit Test by Lafayette Instrument Co, Inc.
2. Bassin Anticipation Timer by Lafayette Instrument Co, Inc.
3. Multi-Limb Coordination by Lafayette Instrument Co, Inc.
4. Multi-Choice Reaction Time Apparatus by Lafayette Instrument Co, Inc.
5. Automated Neuropsychological Assessment Metrics (ANAM) Switching Task (math, spatial orientation, switching)
6. Educational Testing Services—Concealed Words Test (CS-2)
7. Educational Testing Services—Snowy Pictures Test (CS-3)
8. Basic flight knowledge: multiple choice, fill in the blank, and true/false test developed by the researcher
9. Basic aircraft instrument knowledge: multiple choice, fill in the blank, and true/ false test developed by the researcher

NPS Professor Nita Shattuck graciously provided the Lafayette Instruments and ANAM software use license. Educational Testing Services issued a license agreement for the Concealed Words Test (CS-2) and the Snowy Pictures Test (CS-3). The researcher wrote the multiple choice, fill-in-the-blank, and true/false knowledge tests using the 2019 *Federal Aviation Regulations and Aeronautical Information Manual* (FAR/AIM), the 2018 "FAA Airmen Knowledge Test Guide," and the 2016 *Jeppesen Instrument Commercial Book* (Federal Avaiation Administration, 2019; Federal Aviation Administration, 2018; Jeppesen, 2016). Appendix H provides a copy of the tests. Figure 35 shows the set of Lafayette Instrument Co. study instruments.

Figure 35. Lafayette Instrument Co. Equipment

Each BPR name and measurement technique aligned with GSPT requirements, where increased values represented higher capacities. If higher capacity results in smaller values, then the PCE collapses (e.g., faster reaction times result in smaller values). Simply taking the inverse of a measurement or adjusting measurement type ensured GSPT alignment. For example, measuring speed instead of time satisfied the requirement. The following list outlines critical instrument information pertaining to this study. Appendix G contains a series of Tables which document each of the following in detail for each BPR in the study:

- The BPR name
- Name and definition of Fleishman's (1992) ability
- BPR priority ranking per O*NET and Fleishman's *Handbook of Human Abilities* (1992)
- Name of the instrument and its source
- Measurement specifics (adhering to GSPT requirements)
- The exact protocol given to each participant

Some BPRs do not have priority rankings. Primary BPR instruments generated these additional BPRs (i.e., math processing as part of the switching task). The knowledge BPRs support Project III. These tables define HSI domain considerations (resources) as model inputs.

F. PROCEDURES

Chapter III provided overarching study procedures. Project I procedures began with participant consent and demographic data collection. Each participant then completed BPR data collection using a randomized order to avoid potential order effect. Once BPR collection was complete, participants were able to take an optional comfort break. The second phase of the study began with the researcher's orientation to the flight simulator. A short verbal quiz ensured the participants understood the instructions. The first ILS approach (HLT 0) was practice and the researcher provided instruction to each participant. The second ILS approach (HLT 1, Figure 36) included approximately three miles of orientation and instruction. Once the participant reached six miles from the airfield, the researcher ceased instruction. Scoring commenced from mile six to DH at KSEA RWY 34R. Flight performance output data were uploaded from X-Plane 11 to the researcher's computer. The participants attempted landings during each HLT; however, the HLT did not include landing performance. Following the completion of HLT, each participant received a certificate of appreciation concluding the study. Appendix F contains a copy of the certificate.

Figure 36. MBHSI HLT

G. ANALYTICS

1. BPR Data Organization

Participant data collection used paper and pencil for non-software BPR measurements. Appendix D provides the forms generated using Microsoft Excel to capture study data. Analytic formulas in Microsoft Excel automated BPR calculations to minimize risk of errors. Maximum performance for each BPR was transferred to the PCE Calculation

Table (also in Appendix D). The PCE Calculation Table captured capacity data from each participant for transfer to the study analytics Microsoft Excel workbook. The analytics workbook organized all participant capacity and HLTp data in matrix format to be imported by R (R Core Team, 2019) and for further analyses.

2. PCE Analytics

Each participant completed BPR capacity measurements, resulting in linear 19-dimension PCEs. Following data collection, BPR data were normalized in terms of percentile to build the R_A profiles showcased in Project II. These R_A profiles overlay the R_D profiles, also introduced in Project II, to quantify the economic resource relationships between availability and demand at various levels of HLTp. When the vision for MBHSI is fully realized, volumetric PCEs might also be created with non-linear BPR data to account for individual capacity variance (i.e., fatigue, temperature, stress). These data would support the development of human capacity documents at the individual and population levels. Because this study measured capacity once, a linear approach was appropriate.

3. HLTp Analytics

This section provides an overview of the HLTp analytics developed for this project and all subsequent projects. A more detailed presentation of these analytics is provided in Appendices I and J. In this case, safe precision approaches to land a single-engine fixed wing aircraft defined valuable task outcome. Quantitative performance boundary conditions from the FAA defined task measurement. The FAA defines windows of performance deemed safe for aeronautical pilotage during instrument meteorological conditions (IMC). An approach to a runway that most closely adheres to the target, not threshold values across performance metrics (Airspeed, heading, altitude, and CDI), define valuable performance. Arrival to decision height (DH) on-point defines a desired outcome, so a controlled landing or a properly executed diversion can be safely completed. The FAA defines performance windows (that if achieved result in a valuable outcome) in their Instrument Flight Certification Standards (Federal Aviation Administration, 2018a):

1. Airspeed: +/- 10 knots
2. Heading: +/- 10 degrees
3. Altitude: +/- 100 feet
4. CDI: < ¾ deflection

Required heading changes to seek Course deviation indicator (CDI) deflection of 0 led to the removal of heading as a performance metric in the study. The CDI captured Glideslope during the entire approach. This led to the removal of altitude as a performance metric as well. The product of three performance metrics—Airspeed, Course (CDI-H), and Glideslope (CDI-V)—defined the univariate HLTp score. The novice nature of the participants justified a 25 percent performance metric relaxation, the most that could be measured by a full-scale CDI deflection:

1. Airspeed: +/- 12.5 knots
2. Course (CDI-H): < full deflection
3. Glideslope (CDI-V): < full deflection

Data quantity varied among participants due to the dynamic nature of the HLT. Performance bins ensured equal scoring across performance metrics. All three included standardized scoring rubrics for performance errors. Airspeed, Course (CDI-H), and Glideslope (CDI-V) had 10 equal performance error bins. Each Airspeed bin had a range of 1.25 kts for a total performance window of 12.5 kts. Each CDI bin had a range of .25 degree (deg) for a total performance window of 2.5 deg. All text (.txt) data outputs from X-Plane 11 were converted to comma separated value (.csv) files for simple use in Microsoft Excel.

Airspeed:

- Target: 65kts
- Performance window: 52.5kts—77.5kts (25% more than FAA Standards)
- Actual Airspeed captured once per second
- Error in kts calculated and squared

Penalties: assessed if Airspeed exceeded the performance window established. Figure 37 provides an example of the Airspeed indicator.

Figure 37. X-Plane 11 Cessna-172 Airspeed Indicator

Course (CDI-H) and Glideslope (CDI-V):

- Target: 0 deg of deflection
- Performance window: < full deflection (< 2.5 deg or 0—2.499 deg), (25% more than FAA Standards)
- Actual deflection each second listed in decimal degrees
- Error in degrees calculated and squared
- Penalties: assessed if deflection exceeded established performance window.

Figure 38 provides an example of the CDI. The vertical needle informs Course deviation and the horizontal needle informs Glideslope deviation. Each of the white dots indicates .5 deg deflection. In this example the vertical needle indicates the aircraft is left of Course by at least 2.5 deg and is on Glideslope.[18]

[18] While the Glideslope instrument needle is physically horizontal, it is measuring vertical flight deviation. While the Course instrument needle is physically vertical, it is measuring horizontal flight deviation, i.e., how far off Course the aircraft is from the ILS. Therefore, in this study and as illustrated in Figure 38, CDI-H represents Course deviation and CDI-V represents Glideslope deviation.

Figure 38. X-Plane 11 Cessna-172 Course CDI

Univariate Score Derived for Each Participant

The univariate scoring rubric captured the three performance metric scores across the HLT then engaged a multiplicative approach to derive a single HLTp score. The reader is reminded that deviation from perfection is how this HLT is measured, thus an inverse was taken to ensure that as HLTp increases, scores also increase. Additional information is provided in Appendices I and J.

- Multiplicative model *not* additive
- HLTp = (Airspeed*CDI-H*CDI-V)
    - Perfection would result in a score of 1,000,000 (the cube of each of the three sub-tasks valued at 100)
    - Median Performance derived a score of 15,625 per rubric
    - Exponential penalty factor due to the squaring of the error scores
- Lower scores equated to better results, but for the purposes of GSPT, better scores need to be higher. So, inverting the score and multiplying it by 1,000,000 achieved this outcome.
    - Example Excel logic: (1/(PROD(Airspeed, CDI-H, CDI-V)))*1,000,000
- Penalties
    - HLTp w/Penalty = (Airspeed w/Penalty*CDI-H w/Penalty*CDI-V w/Penalty)
- Performance Thresholds
    - <= 10 penalties for each sub-HLTp AND <=25 for HLTp (PASS or FAIL)

The HLT value desires zero deviation at all performance assessments during the task. A score of 1,000,000 is where zero deviations occur at all performance assessments during the entire HLT, across all three performance metrics. A product score of 15,625 defines median performance values across all three metrics for all assessments:

$$\text{Median performance deviation multiplier} = 5$$

therefore,

$$5^2 = 25$$

and

$$25 * 25 * 25 = 25^3 = 15{,}625.$$

This approach incentivized performance and penalized deviation using an exponential scoring rubric. It also provided high diagnostic capacity during performance analytics. The novice nature of the participants led to lower than median scores; many scored 0. This confirmed the HLT to be complex and to have a high demand for cognitive and psychomotor resources. The participants were informed of the scoring method, which emphasized an exponential increase in penalty for larger target deviations.

H. FORECAST MODEL FRAMEWORK

HLTp data analytics were captured in Microsoft Excel. The data output function in X-Plane 11 supplied all HLTp data. Data collection included speed, CDI-H, CDI-V, altitude, distance from airfield using Distance Measuring Equipment (DME), and Global Position System (GPS) coordinates to assist in trimming non-HLTp data. The software collected performance data at a rate of 1Hz, every second. Microsoft Excel imported the data as .txt files then converted them to .csv files. The net HLTp data set for each person included five miles of the ILS approach using DME measurements. Three columns included the critical Airspeed, Course (CDI-H), and Glideslope (CDI-V) data. Absolute values converted negative CDI data to ensure an all-positive value data set. Table 6 illustrates a sample of data prepared for entry into the analytic worksheets. Data outside the five-mile study section of the approach were discarded.

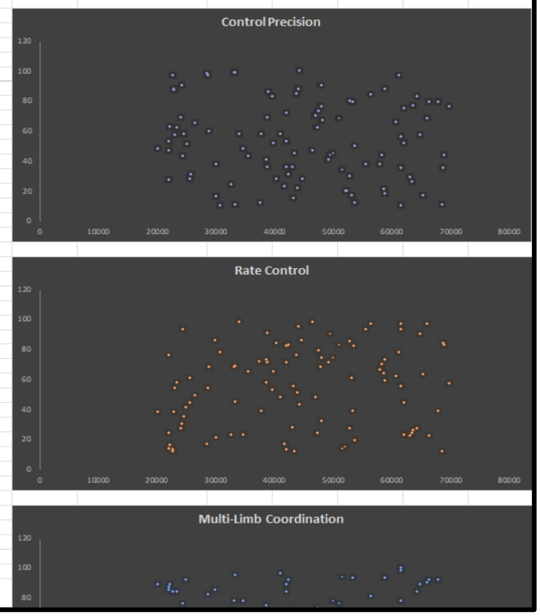
Table 6. Preparation of Translated HLTp Data for Entry into Analytic Worksheets

KIAS	__alt,ftmsl	NAV_1,h-def	NAV_1,v-def	__DME,_dist		KIAS	H-Def	V-Def
68.66396	2062.13184	-0.28547	-0.32154	6.01813		68.66	0.285	0.322
68.40506	2066.25659	-0.29697	-0.27703	5.99886		68.41	0.297	0.277
66.98376	2074.28711	-0.31013	-0.20374	5.97993		66.98	0.310	0.204
65.2513	2081.86401	-0.32398	-0.12844	5.96138		65.25	0.324	0.128
63.70247	2088.07080	-0.33881	-0.06048	5.94320		63.70	0.339	0.060
62.27673	2094.77612	-0.35521	0.00678	5.92530		62.28	0.355	0.007
60.82193	2101.25000	-0.37225	0.07587	5.90755		60.82	0.372	0.076
59.66214	2105.91846	-0.38947	0.13519	5.89028		59.66	0.389	0.135
58.79191	2109.65894	-0.40719	0.18803	5.87301		58.79	0.407	0.188
58.1106	2112.71899	-0.42519	0.23691	5.85590		58.11	0.425	0.237
57.59738	2115.39087	-0.44375	0.28271	5.83906		57.60	0.444	0.283
57.3217	2117.63062	-0.46279	0.32709	5.82220		57.32	0.463	0.327
57.42306	2119.33203	-0.48244	0.36811	5.80511		57.42	0.482	0.368
57.70288	2121.34546	-0.50156	0.41046	5.78803		57.70	0.502	0.410
58.08062	2122.80981	-0.52102	0.45192	5.77090		58.08	0.521	0.452
58.72458	2122.66943	-0.54065	0.48537	5.75365		58.72	0.541	0.485
59.77372	2120.12793	-0.56001	0.50644	5.73604		59.77	0.560	0.506
61.07315	2117.34180	-0.57934	0.52303	5.71800		61.07	0.579	0.523
62.2491	2115.50952	-0.59759	0.54498	5.69972		62.25	0.598	0.545
63.26049	2113.72998	-0.61521	0.57015	5.68114		63.26	0.615	0.570
64.34718	2110.72339	-0.63167	0.58901	5.66251		64.35	0.632	0.589
65.44327	2108.50195	-0.64667	0.61026	5.64322		65.44	0.647	0.610
66.18784	2108.36890	-0.65981	0.64312	5.62403		66.19	0.660	0.643

The study analytics workbook collected participant BPR and HLTp data as inputs for the appropriate scatterplot models. All MBHSI model development was done using Microsoft Excel and R statistical programming language (R Core Team, 2019). The data were entered into R for RDF development with ease, as Project II demonstrates. A simplification using randomized data (Table 7) illustrates this method in Microsoft Excel. The participant number is randomly generated. There is underlying code in the worksheet that takes (in this case) 6 of the 19 BPRs, calculates the HLTp, and plots each participant's data for the BPRs. Note the even distribution of the random data in the plots, then contrast them with actual MBHSI data plots shown in the results section.

Table 7. Example Forecast Model Framework

Participant	BPR Score 1	2	3	4	5	6	HLTp
1122	77	26	50	44	15	70	63643
2871	24	23	72	90	94	21	32663
3061	83	53	52	39	38	24	39615
3887	17	61	35	66	27	35	53272
3262	68	83	76	35	12	57	51097
9182	70	48	43	85	13	93	46983
3786	45	12	72	12	31	29	43449
9823	90	30	47	11	16	30	24309
5300	62	24	73	100	72	88	47374
835	73	79	51	95	84	62	47443
7905	99	68	78	99	70	87	33109
9379	48	38	89	42	23	25	20130
3228	58	48	96	71	25	38	41051
2266	51	41	92	44	71	37	25065
6204	17	63	64	40	11	14	65479
9811	43	93	76	24	15	10	24486
9923	18	73	93	97	89	62	58903
9848	86	91	62	44	14	30	38866
6433	44	70	17	91	82	49	58402
3817	100	43	59	42	68	79	44173
2236	79	39	92	63	87	80	67904
8767	68	97	90	81	34	93	66134
4586	22	51	15	45	92	17	43867
3538	53	82	41	51	73	22	41989
4163	50	82	55	77	92	87	53709
3548	88	95	34	85	95	83	44014
3642	38	93	16	22	50	80	55733
9258	48	23	78	28	99	67	34746
8337	58	35	26	93	62	21	24602
4887	38	21	19	69	37	79	30047
3601	57	54	31	91	69	100	23111
58	43	65	61	16	12	61	35512
5288	58	39	45	28	92	70	37735
8923	35	84	19	49	47	42	68757
3445	41	71	50	66	77	54	49121
1855	36	28	59	79	58	89	43121
2772	76	57	28	84	92	65	69869
9276	69	71	23	40	23	42	38794
2729	66	62	21	37	47	55	60748
6897	99	45	95	67	51	31	33253



I. RESULTS

Overall, most of the models demonstrated desired GSPT characteristics. Scores ranged from 0 to 14,015 (*Mdn* = 282) for HLT 1 without penalties enforced and (*Mdn* = 111) for HLT 1 with penalties enforced. A total of 19 BPRs were measured and BPR data from the knowledge tests were withheld from model building. Representative GSPT models were not generated for these tests due to the nature of multiple-choice knowledge tests and the novice participants. The participants who received training and took the posttest provided BPR data that was important for Project III analytics. Therefore, 17 BPRs remained for model construction. The 17 separate BPR data (inputs) for 65 participants (64 plus a cadaver point) and HLTp data (outputs) supported model development.[19] While the study collected data for two HLTs (HLT 1 and HLT 2), the following data from each participant-built Project I models:

[19] During Project II RDF development, it became clear that a point at the origin (0,0) would assist in grounding the RDFs to the origin. At the recommendation of Dr. George Kondraske, a "cadaver" participant was added, zero BPRs and zero HLTp. This point is referenced as participant 65 in this study. This issue is discussed further in Chapter V.

- Capacity data across 19 BPRs
- HLTp 1
- HLTp 1 w/penalties enforced
- Airspeed sub-HLTp 1
- Airspeed sub-HLTp 1 w/penalties enforced
- Course (CDI-H) sub-HLTp 1
- Course (CDI-H) sub-HLTp 1 w/penalties enforced
- Glideslope (CDI-V) sub-HLTp 1
- Glideslope (CDI-V) sub-HLTp 1 w/penalties enforced

In total, the data spreadsheet consisted of 65 rows (participants) and 90 columns of data for a total of 5,850 data points. Each of the two HLTs generated an estimated 2,500 .txt data points from X-Plane 11 for each participant that were trimmed to approximately 825, then converted to .csv, and imported into Microsoft Excel. Each participant HLTp (1 and 2) contained approximately 1,650 data points for a total of 105,600 HLTp data points in the study. Combined with BPR data, this study consists of over 111,000 separate data points.

Figure 39 illustrates the distribution of HLTp 1 w/o penalty scores. These data speak to the complexity and challenge of the task. The median (*Mdn* = 282) for HLTp 1 w/o penalties enforced is closest to participant 31 on the x-axis.

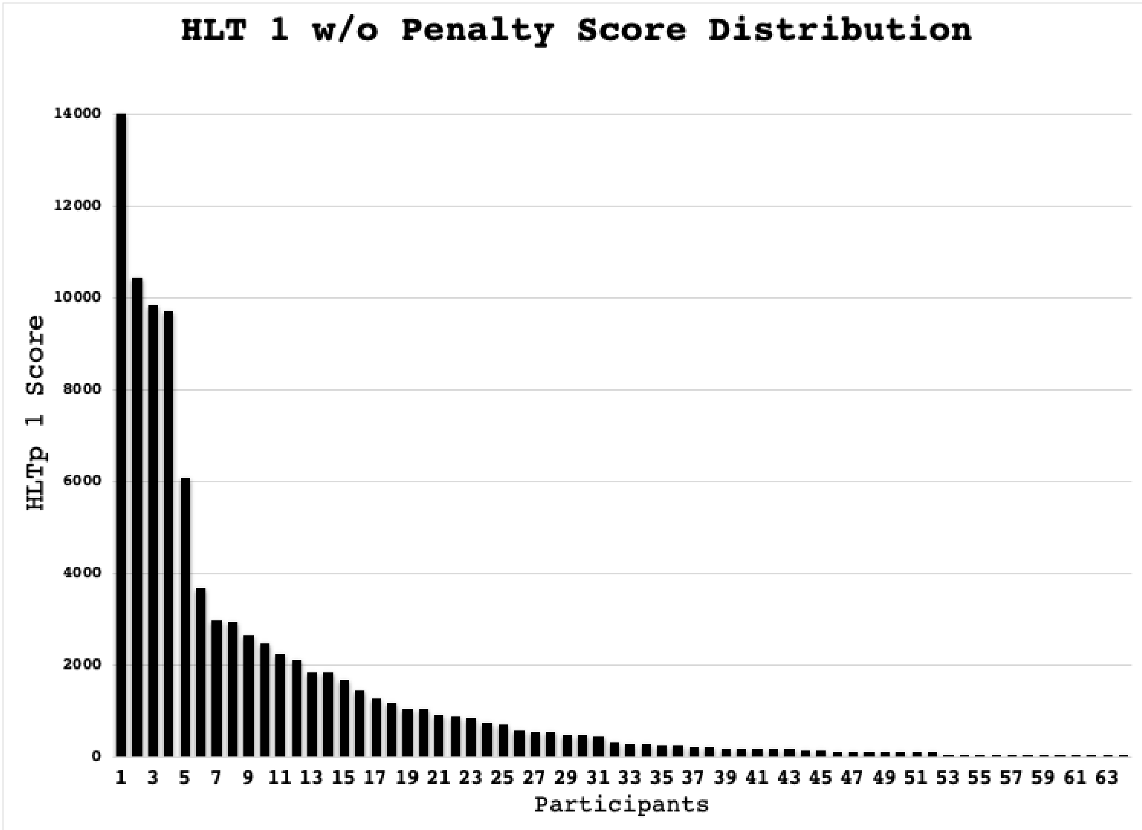


Figure 39. HLTp 1 with Penalties Enforced Scores

The GSPT methodology and this study's HLT resulted in predominately consistent scatterplots with GSPT characteristics detailed in Chapter II and illustrated in Figure 40. In this study, as will be the case during operational MBHSI efforts, quality models serve as the foundation for subsequent MBHSI work documented in Projects II-III. While this study contains adequate data to generate hundreds of GSPT models, 136 were built for Project I to facilitate Projects II-III of this laboratory study. Nearly all (approx. 130/136) MBHSI models reflect excellent GSPT characteristics depicted in Figure 41. Those that are not excellent demonstrate good characteristics. Examples of these scatterplots are discussed later in this results section. These few models could easily be considered excellent if a few data points are categorized as outliers. Each model also demonstrates unique, non-linear, lower threshold boundaries. An illustrative sample of models is provided in this section; however, the complete Project I model set of 136 scatterplots and their numerical identifiers are provided in Appendix G. The appendix presents 17 scatterplots (one for each BPR) for each of the following eight datasets:

- BPR (1-17) and HLTp 1
- BPR (1-17) and HLTp 1 w/penalties
- BPR (1-17) and Airspeed sub-HLTp 1
- BPR (1-17) and Airspeed sub-HLTp 1 w/penalties
- BPR (1-17) and Course (CDI-H) sub-HLTp 1
- BPR (1-17) and Course (CDI-H) sub-HLTp 1 w/penalties
- BPR (1-17) and Glideslope (CDI-V) sub-HLTp 1
- BPR (1-17) and Glideslope (CDI-V) sub-HLTp 1 w/penalties

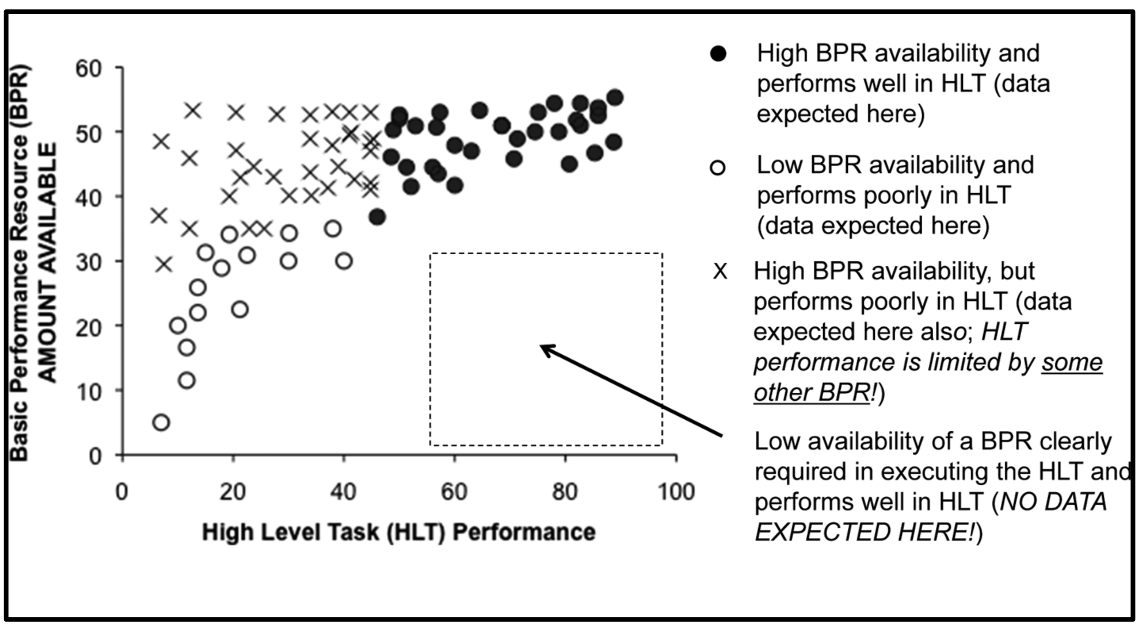


Figure 40. Target GSPT Model Characteristics.
Source: Kondraske (2011, p. 249).

Figure 41 is an excellent representative sample from the study. This HLT 1 model reflects target GSPT characteristics articulated in Figure 41. A few key characteristics are worth noting. First, the high-density of data near zero on the x-axis indicates the HLT was very difficult for most participants. Second, the red line in Figure 41 highlights the increase resource demand with higher HLTp values. Third, many participants demonstrated excellent capacity for this particular BPR (visual motor tracking accuracy), yet performed poorly during the HLT. This suggests different BPRs limit their HLTp. Finally, no participants that demonstrated low BPR capacity performed well on the HLT. This evidence suggests this is an excellent BPR candidate for NCRA performance forecasting.

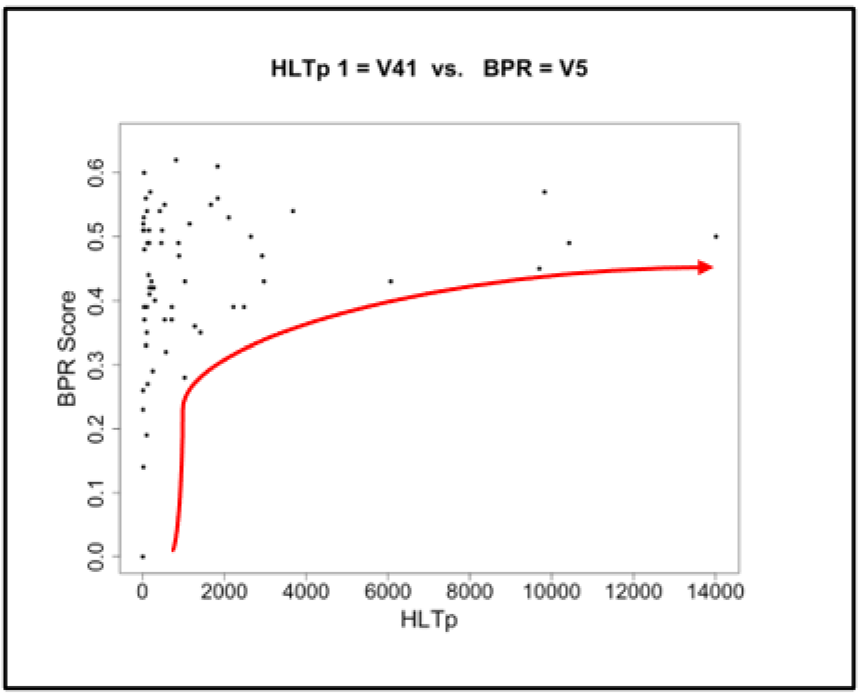


Figure 41. Representative Sample from HLTp 1 Model Set (BPR 5 = Visual Motor Tracking Accuracy)

Figure 42 is the same model with penalties enforced and demonstrates score reduction, predominately in the lower-scoring participants. Higher-scoring participant data remain stable with penalty enforcement because, in general, those participants' scores remained in the performance windows. The reduction in data along the lower output values provides lower resolution for RDF development. Note the increase in scores of 0 in Figure 42. Penalty enforcement may help reduce Type I errors (false positives) in personnel selection, for example. Project II examines Type I and II (false negatives) performance forecast errors in more detail.

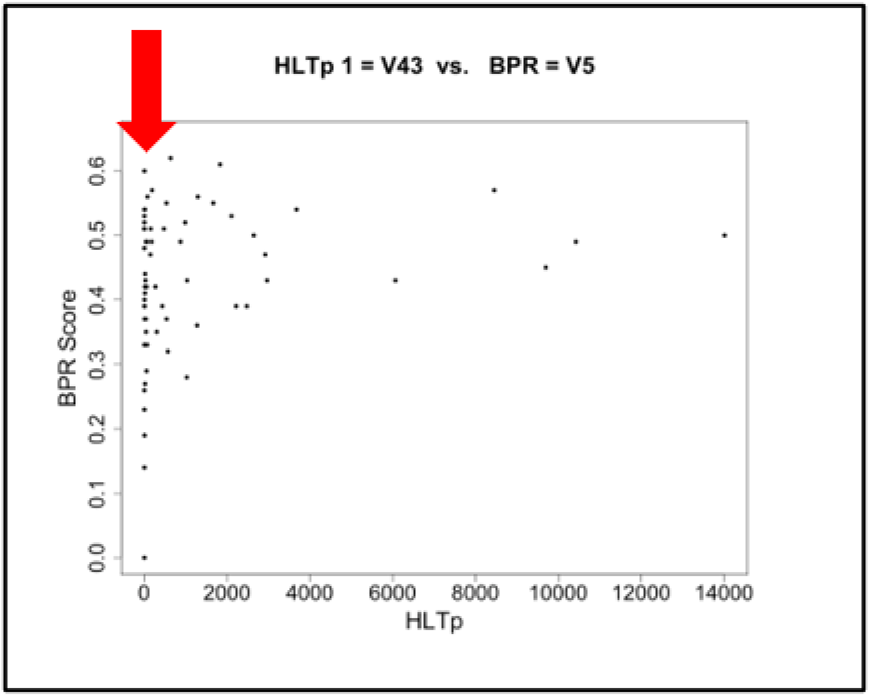


Figure 42. Representative Sample from HLTp 1 (Showing Impact of Penalties Enforced) Model Set (BPR 5 = Visual Motor Tracking Accuracy)

Figure 43 illustrates a representative sample from the Course sub-HLTp 1. A few key characteristics warrant a brief discussion. First, an explicitly clear threshold envelope can be detected just above the red line. This BPR (math speed) demonstrates a very linear relationship with HLTp. Second, no participants demonstrated low capacity and high HLTp scores. Third, the wider dispersion of data across the HLTp axis suggests this sub-task was not exceedingly difficult. Because the simulated ILS approach was conducted under VFR conditions, this finding is not surprising.

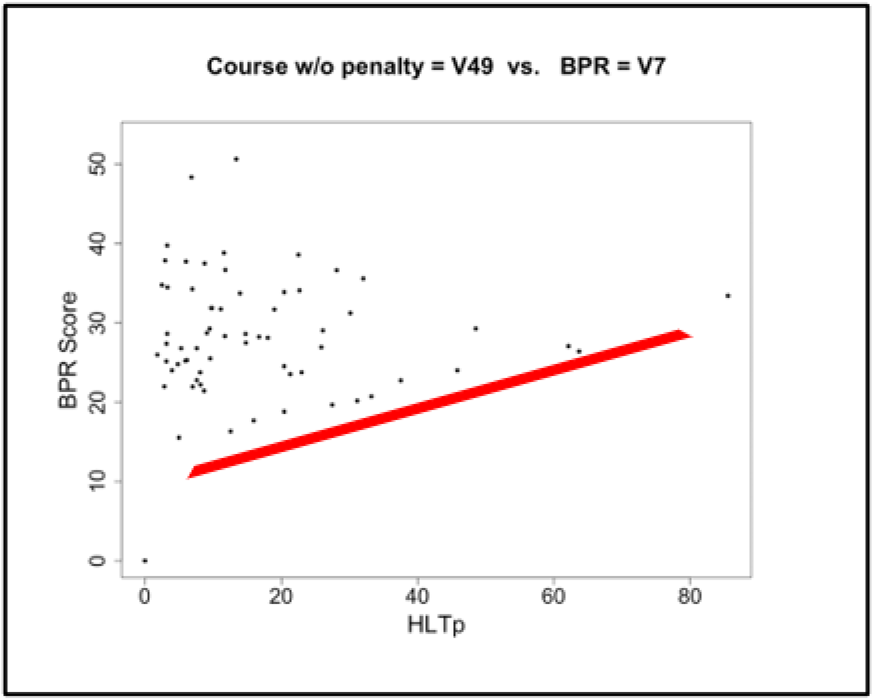


Figure 43. Representative Threshold Sample from Course Sub-HLTp 1 Model Set (BPR 7 = ANAM Math Speed)

The ceiling effect is apparent in some of the accuracy BPRs, such as the Multi-Limb Coordination Accuracy Test. This interesting phenomenon might indicate that participants focused too heavily on accuracy rather than on speed, see Figure 44.[20] Ceiling effect data can contribute to under- and over-prediction when performance is forecasted because the RDF ascends vertically then immediately pivots horizontally. Therefore, when forecasting performance, the BPR value seeking the RDF may prematurely contact the RDF resulting in a lower than observed HLTp or it may never contact the RDF resulting in extreme over prediction. The practitioner is cautioned not to remove these BPRs in the event the stakeholder prioritizes accuracy as part of HLT. This study assumed accuracy was a priority, thus these models remain in Project II. A potential solution for the ceiling effect may include having participants continue the test with increasing speeds until errors in accuracy arise. Perhaps participants could have completed the BPR test faster than

[20] Future MBHSI BPR data collection should consider having participants re-attempt the task until errors are noted to avoid ceiling effect in models.

observed, given Figure 44. The observation of a lesser ceiling effect produced by an ANAM Accuracy Test in Figure 45 provides a contrast of accuracy BPRs. A simple adjustment in protocol and instruction to the participants may alleviate this issue. In an effort to maintain protocol integrity, nothing changed during data collection once this phenomenon was evident.

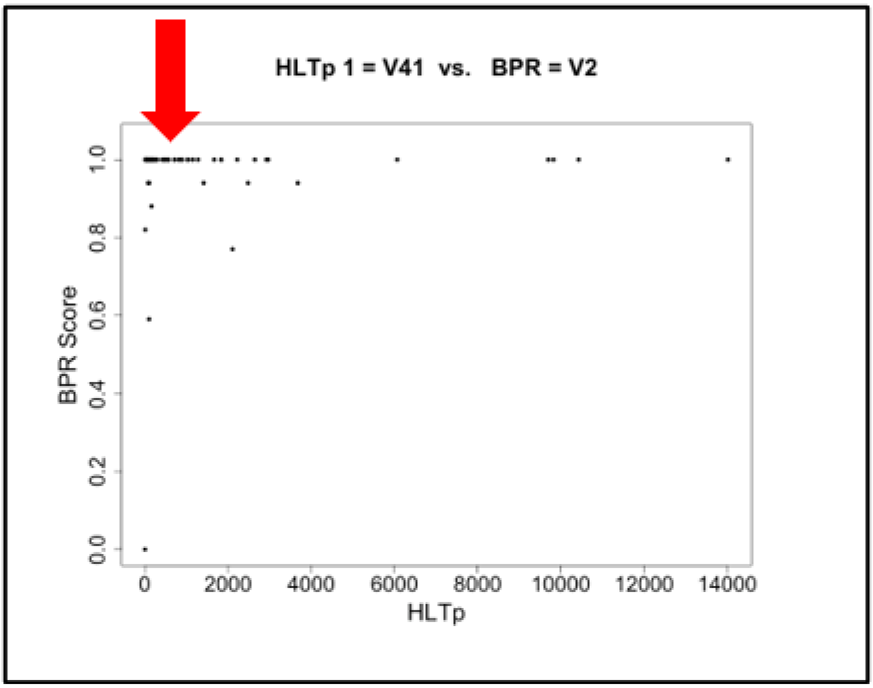


Figure 44. Example of GSPT Ceiling Effect (BPR 2 = Multi-Limb Coordination Accuracy)

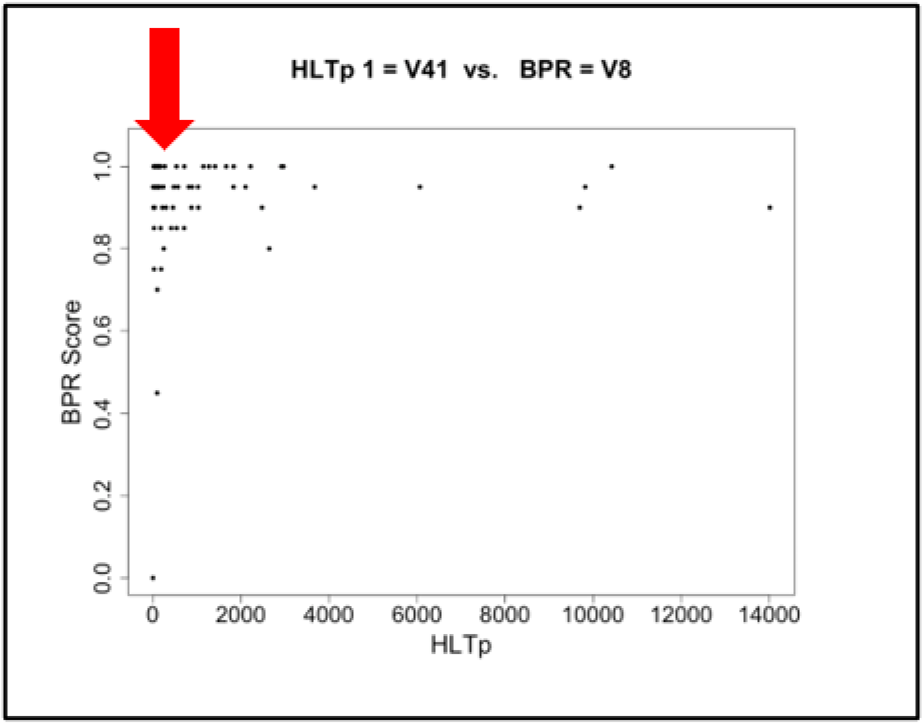


Figure 45. Example #2 of Lesser Ceiling Effect (BPR 8 = ANAM Math Accuracy)

In review of the data collected in accordance with GSPT, the scatterplots reflect GSPT characteristics. The reader is encouraged to review the entire dataset in Appendix M. Of the entire the dataset, only two BPRs were withheld from Project II: the two knowledge tests (V18 and V20). These models do not demonstrate strong GSPT characteristics. Figure 46 illustrates their weak characteristics. Specifically, the data demonstrate a normal distribution across the BPR y-axis (i.e., most scored mid-range with fewer scoring high and low). These results likely are due to novice participants taking a primarily multiple-choice test. Knowledge BPR data demonstrate value in Project III pre-posttest analytics. Future knowledge BPR considerations are also described in Project III.

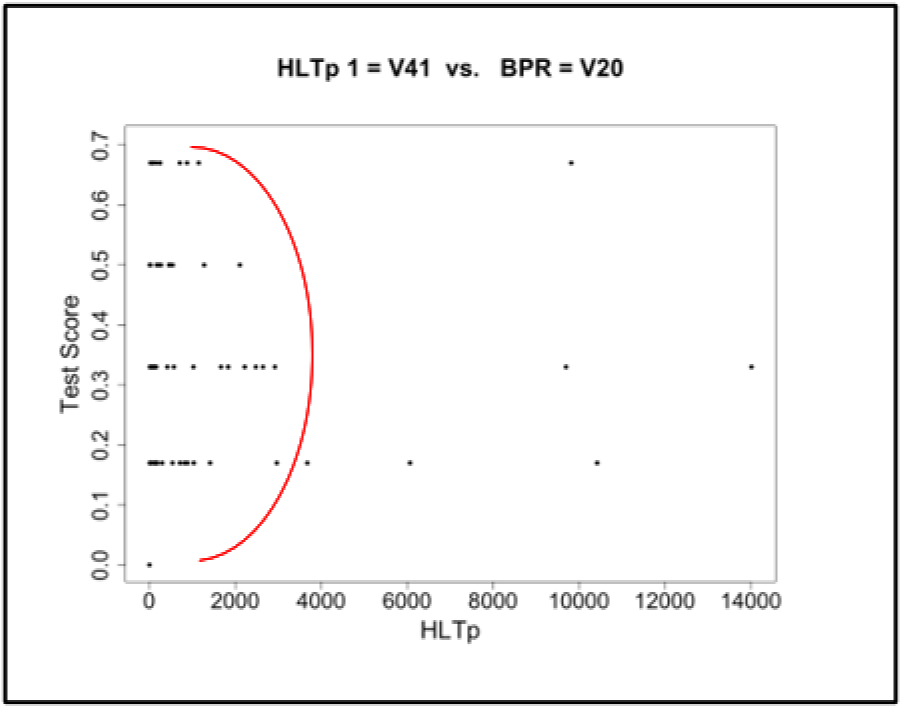


Figure 46. Example of Near Zero BPR and Chance (Exhibiting Normal Distribution)

Figure 47 illustrates perhaps the weakest model in the study. The two data points occupying the lower right corner of the model violates GSPT preferred characteristics. However, if these two data points are removed as outliers, a clear threshold becomes evident. These two points could be investigated in detail regarding the BPR measurement (the concealed word test). In this case, these participants simply did not demonstrate capacity for an unknown reason during data collection. Consistent with typical human performance research, addressing outliers presents distinct challenges. MBHSI will need to investigate appropriate methods for identifying and responding to potential outliers.

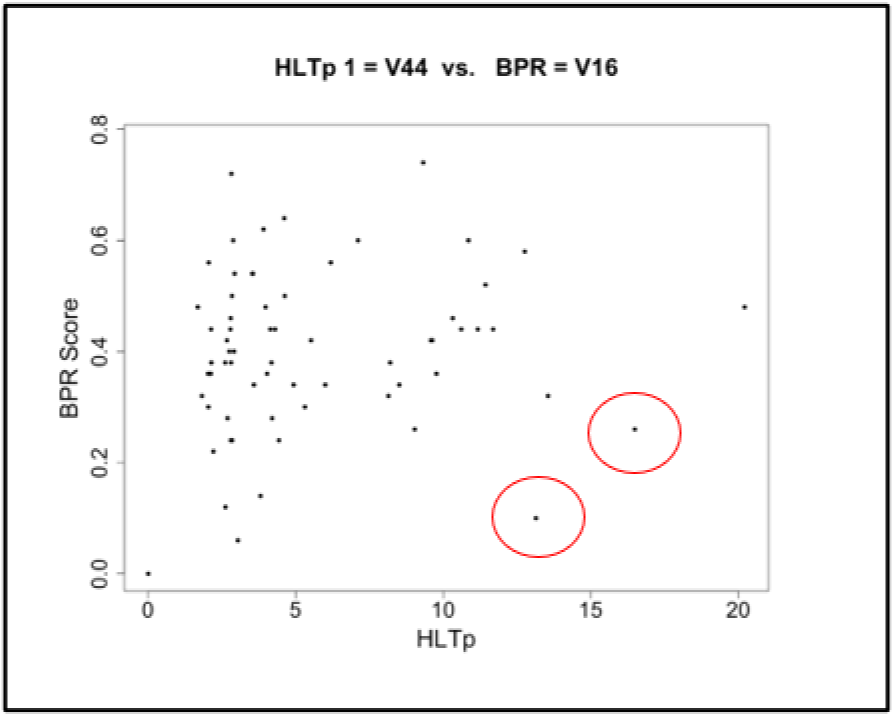


Figure 47. Example of Weaker Model (Airspeed Sub-HLTp 1 vs. BPR 16 = Perceptual Integration Capacity—Concealed Words)

J. IN CLOSING

The first stated purpose of MBHSI performance forecast model development was to realize the outputs of GSPT as a contained system within MBHSI. The establishment of model input and output relationships defined the requirement of GSPT. The second purpose was to plan, execute, and document MBHSI alignment with the first HSI functional requirement, *define HSI domain considerations (resources) as inputs*. The results of MBHSI Performance Forecast Model Development, a series of scatterplots, agree with the expectations detailed by Kondraske (2011). The plots represent a visual and quantitative relationship between human resources (BPRs) and overall system performance (HLTp) as hypothesized. Therefore, Project I's stated purposes were achieved. The models produced using these methods serve as prerequisites for fitting threshold envelopes (RDFs) as described by NCRA in Chapter III. Project II establishes a novel method for quantifying RDFs and forecasting system performance.

V. RESEARCH PROJECT II: RESOURCE DEMAND FUNCTIONS AND SYSTEM PERFORMANCE FORECASTS

> Consumers of forecasting will stop being gulled by pundits with good stories and start asking pundits how their past predictions fared—and reject answers that consist of nothing but anecdotes and credentials.
>
> —Philip E. Tetlock, *Superforecasting*

A. INTRODUCTION

1. Overview

The first purpose of this study was to investigate the conceptual outputs of NCRA to enable MBHSI. Satisfying this requirement enables Model-Based Systems Engineering (MBSE) as shown in Figure 48. This improved capacity of HSI suggests a complementary model-based approach to MBSE may be possible.

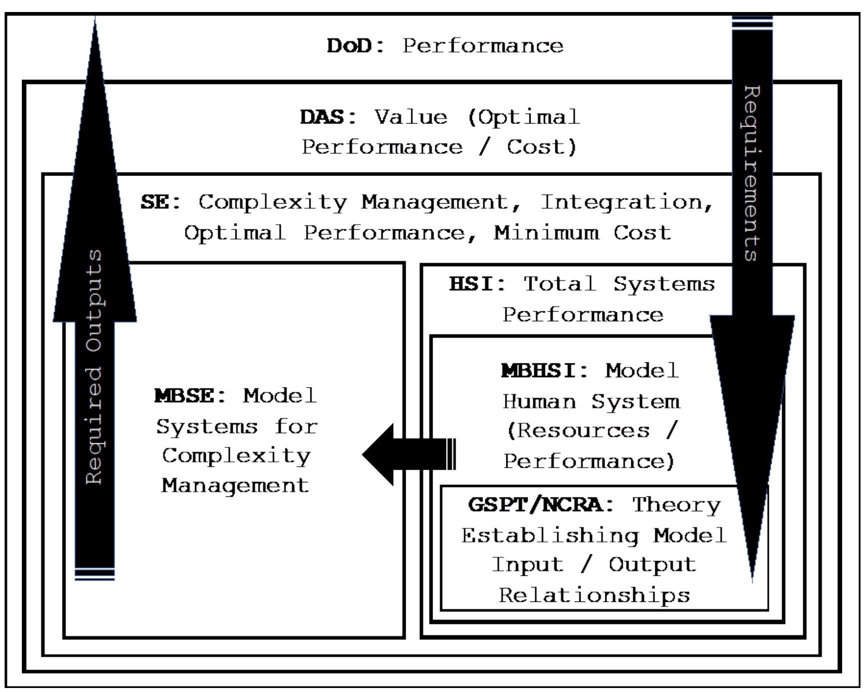


Figure 48. MBHSI Requirements Hierarchy and Related Outputs. Adapted from Hitchens (1992, p. 10).

The second purpose of this study was to satisfy the second, third, and fourth MBHSI Functional Requirements (MBHSI-FRs), which state to:

- establish a quantitative relationship between HSI domain resources and system performance
- define a model that delivers performance data as outputs
- measure performance forecast accuracy.

Figure 49 illustrates the serial MBHSI-FRs as mapped to this study.

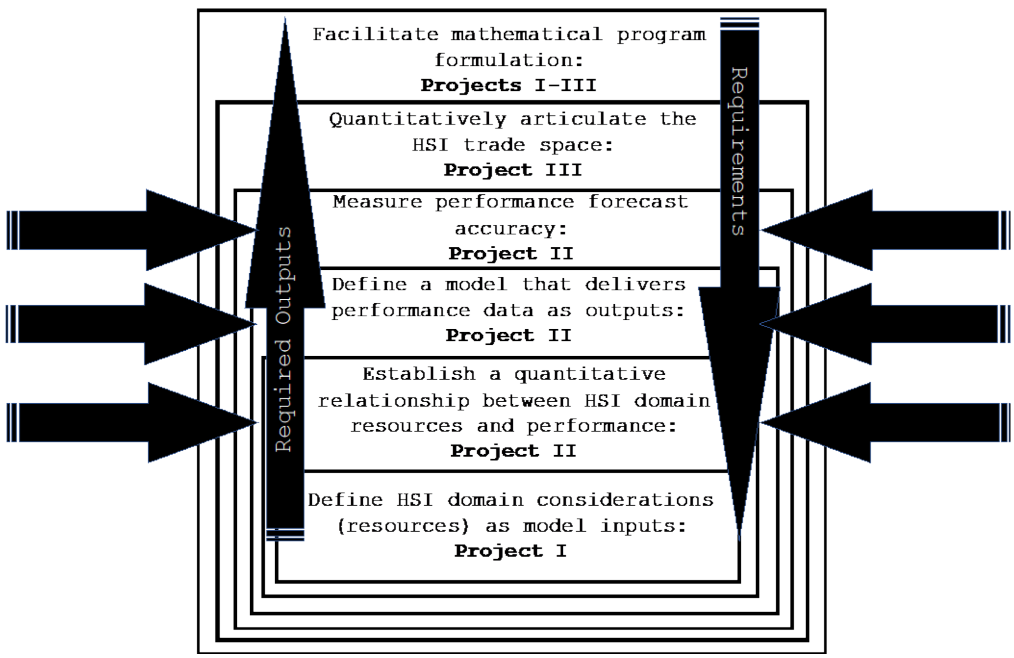


Figure 49. Project II as Mapped to MBHSI-FRs

The goal of this study was to quantify relationships between the resources (inputs) and system performance (outputs) so that either could be forecasted. This explicit relationship, the resource demand function (RDF) establishes this forecast capacity for MBHSI. This capacity requires the development of a threshold-based envelope within the models, the MBHSI RDF. This development requires the use of Project I models, the concepts of NCRA, and the novel MBHSI method to fit the RDFs. Median forecast accuracy and actual agreement between observed and predicted values validate the models. Agreement between forecasted and observed performance suggests resource available (R_A) and resource demand (R_D) profiles can be reliably constructed. These profiles establish observable evidence of improved HSI capacity to enable SE. A roadmap for this Project II illustrates this process in Figure 50. The reader will complete four laps around the roadmap in this study as the HLT and the three sub-HLTs unfold in an orderly fashion.

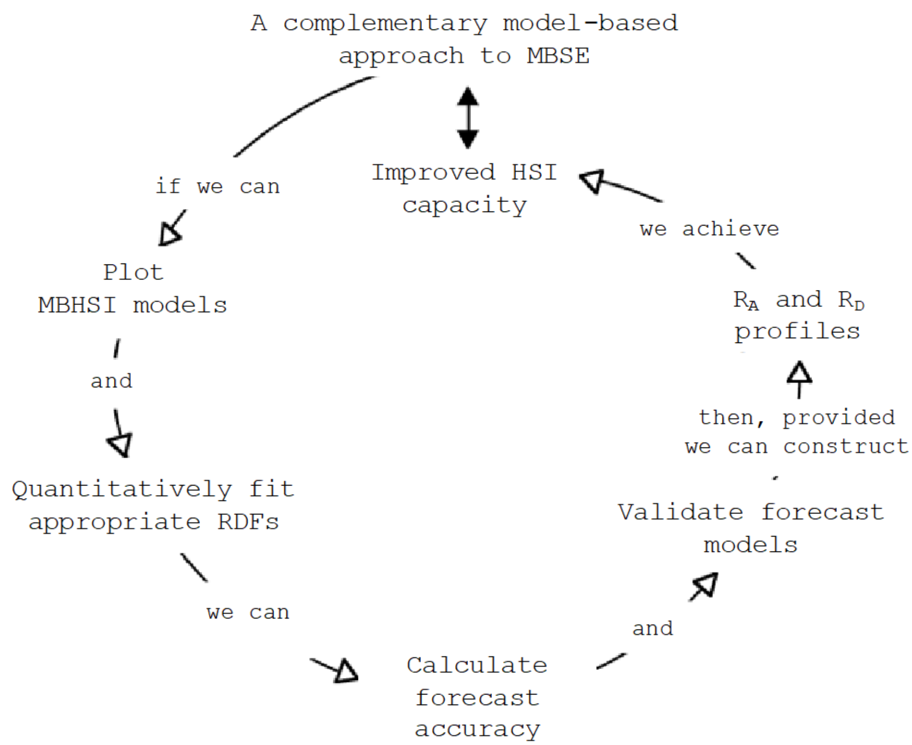


Figure 50. Project II Roadmap. Adapted from Hitchens (1992).

2. Problem Statement

The practice of HSI involves forecasting future system performance, given system inputs (human considerations and technological design decisions) that derive outputs (total system performance and cost) (Tvaryanas, 2010). Department of Defense (DOD) HSI practitioners do not have a holistic and validated method to forecast accuracy. As a contributing discipline to SE, HSI lacks a complementary model-based approach to integrate the human system into the MBSE system architectures. To support MBSE, HSI must develop complementary methods for modeling, forecasting, executing trades, and optimizing performance. This project demonstrates a theoretically based process to construct accurate RDFs and system performance forecasts. These RDFs and forecasts serve as the critical prerequisites for constructing valuable resource profiles. The insight provided by the resource profiles may increase the capacity for HSI to address trade space and optimization challenges.

3. Project II Objective

The objective of this study was to establish the RDFs to generate system performance forecasts. A novel approach to NCRA served to quantitatively fit the RDFs to Project I models. Outputs from the RDFs in this study support the remaining two research projects in this dissertation.

4. Hypothesis

NCRA can reliably forecast system performance by quantitatively modeling the relationship between system inputs (BPRs) and outputs (HLTp).

5. Methodology

This study established the quantitative relationships between R_A and HLTp by applying the concepts of NCRA to the Project I GSPT models. RDFs were fit to each Project I model using a unique form of quantile regression (quantile = .07) and novel code logic developed in R. Each set of HLTp and sub-HLTp models included all 17 BPRs. A jackknifing technique validated each set of models using a 'leave one out, fit, replace' process for all participants in each model (17 BPRs x (65-1 participants) = 1,088 models). The cadaver point was included during RDF fitting, but not validation as it has zero R_A and HLTp. A quantile optimizer evaluated each model for minimum difference between forecasted and observed performance using 15 different quantile values (17 BPRs x (65-1 participants) x 15 quantile levels = 16,320 models). The models were appraised for accuracy by comparing median and individual forecasted HLTp to observed HLTp. Manual RDF adjustments demonstrate capacity to target accuracy or Type I or II error avoidance. Limiting BPRs by study sample set and individual were identified. Individual R_A and R_D profiles generated quantitative insights regarding their relationships at various HLTp values. The study results established initial MBHSI forecast accuracy levels and support Project III: MBHSI Trade Space.

6. Sample

Project I models supplied the study sample set. Data from the 64 participants and the one "cadaver" point totaled a sample size of 65. This point at the origin of each model grounded the MBHSI RDFs, improving model accuracy. Specifically, this point has zero BPRs and zero HLTp. Periodic reviews of Project I models during data collection ensured adequate resolution for RDF fitting at 50 participants. In the operational environment, stakeholder-determined HLTp will provide an initial zone along the HLTp scale to target RDF resolution. This study did not have that constraint.

A complete list of BPR and performance vectors can be found in Appendix G. In an effort to communicate and document the study results efficiently, the following model sets demonstrate a representative sample of Project I models for this study's analytical purposes:

- BPR (1-17) and HLTp 1
- BPR (1-17) and Airspeed sub-HLTp 1
- BPR (1-17) and Course (CDI-H) sub-HLTp 1
- BPR (1-17) and Glideslope (CDI-V) sub-HLTp 1.

7. Research Design and Variables

The study was a within-subjects pretest measurement model-building design. The laboratory study outputs were HLTp and inputs were BPRs. The specific conditional expectation is:

$$E\,[HLT_P|R_A].$$

A unique aspect of NCRA modeling theory suggests the presence of a second forecast conditional expectation. The first conditional expectation builds the RDF. Once the RDFs are established, a reverse model forecasts R_A (BPR) requirements for a given value of HLTp. The specific conditional expectation is:

$$E\ [R_A|HLT_P].$$

These two models generate both performance forecasts and MBHSI R_D profiles shown in the results section (Chapter V, Section B) of this study. Figure 51 illustrates the two conditional expectations described.

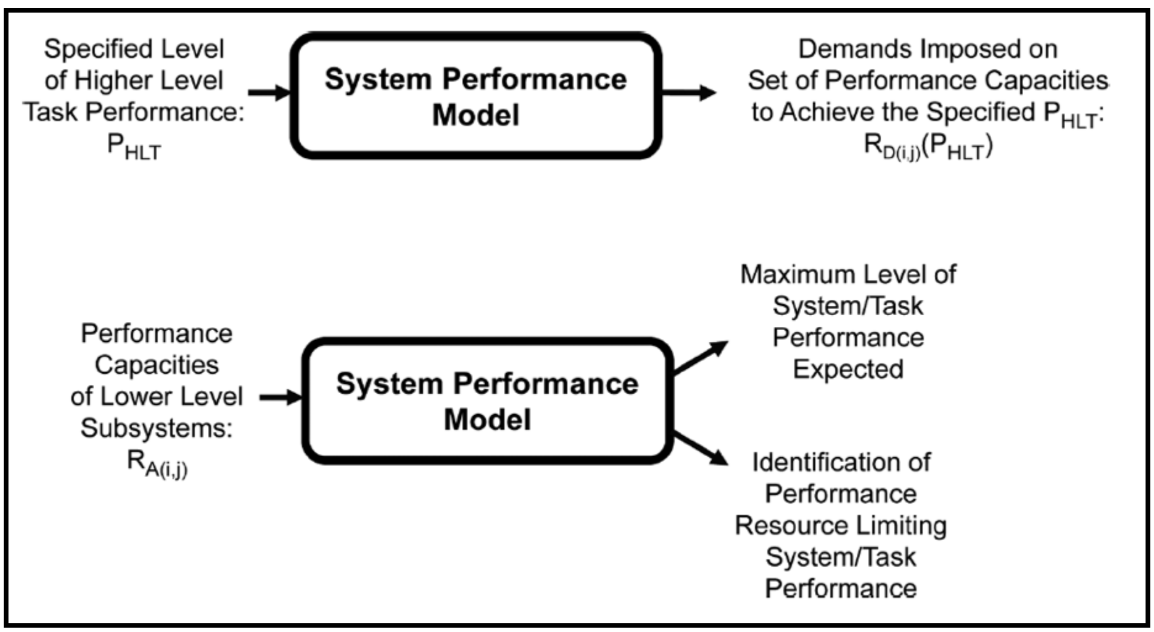


Figure 51. GSPT/NCRA System Performance Model. Source: Kondraske (2011, p. 237).

8. Instruments

The study utilized Microsoft Excel and R statistical programming language. Excel organized and stored the study's data set as described in Project I. R provided the coding language and analytical capacity to generate the novel MBHSI RDF functions. Over 500 lines of R code represent the various functions and analytical tools used in this study. For patent reasons, the dissertation does not present the code logic.

9. Procedures

Development of the RDFs and system performance forecasts adhered to the following procedures. In the results section (Chapter V, Section B), the outputs from steps 3–5 are documented in step 6. This is because multiple iterations of steps 3–5 are required to arrive at appropriate RDF fits.

1. Plot MBHSI models and examine for errors (from Project I).
2. Construct RDFs along each BPR in the model set.
3. Validate models using the jackknifing technique.
4. Execute optimization technique to evaluate quantile (tau) values minimizing the difference between actual and predicted HLTp scores across all participants.
5. Execute manual adjustment technique to find tau values that seek minimum differences between actual and predicted HLTp scores across all participants.
6. Appraise the difference between average predicted and actual HLTp values.
7. Generate pareto data for limiting BPRs across the HLT.
8. Construct participant RA profiles.
9. Construct RD profile for a set of HLTp values.

The study analyzed the four HLTp 1 model sets specified in Chapter V, Section A.1. Three of these model sets demonstrated the ability to move down the hierarchy of the HLT by examining the subtasks of Airspeed, Course (CDI-H), and Glideslope (CDI-V). Additionally, during HLTp 1 analysis, two separate HLTp values demonstrate the utility and capacity of MBHSI RDFs to align with stakeholder error prevention and/or accuracy priorities (i.e., different RDF fits can maximize accuracy or minimize Type I or Type II errors). For example, if Type II error avoidance is a priority for the stakeholder, then the MBHSI RDF can be set to reduce false negatives. However, accuracy tradeoffs may exist if quantile values are adjusted. These tradeoffs between error avoidance and accuracy can be quantified by comparing outcome differences. Performance forecast accuracy was evaluated using an HLTp score of 282, the observed median value. An HLTp of 1,500 was also evaluated to demonstrate the explicit accuracy tradeoffs mentioned.

10. Analytics

Data analysis and RDF creation was performed using R with code developed specifically for this dissertation. Model validation used the jackknifing technique to compare median actual and predicted values. Specifically, each participant HLTp forecast was executed using a model built with the remaining 63 participant data sets. Table 8 illustrates an example of quantile optimization results to include optimal mean and median quantile values. The optimizer fit 15 different RDFs for each BPR then compared forecasted HLTp scores to observed HLTp scores. Following this process, the quantile that resulted in the lowest difference between forecasted and observed HLTp values was returned. For example, the optimizer identified a quantile value of .09 for V1 (BPR1 and HLTp 1) as the most accurate quantile for this particular model. Mean and median values are provided across the BPR set in the table. These values should be investigated by the practitioner for fit during steps 3–5. There is not likely anything significant regarding the same value being returned for both mean and median in this example. This level of analytic capacity suggests that each model RDF can be optimally fit in accordance with stakeholder priorities (e.g., error avoidance, accuracy, etc.). This study used a single quantile value (quantile = .07) for the set of models analyzed because this value appeared to generate the most accurate forecasts across the sample set. If Type II error avoidance was desired, a quantile of .05 or .03 might be investigated. The quantile determines how tight the RDF is fit to the scatterplots.

Table 8. Optimization Results for Quantile (tau)

MBHSI RDF Quantile (*tau*) Optimizer Results by BPR (Range .01 – .15)									
V1	V2	V3	V4	V5	V6	V7	V8	V9	Mean *tau*
0.09	0.06	0.10	0.08	0.09	0.10	0.09	0.06	0.09	0.09
V10	V11	V12	V13	V14	V15	V16	V17		Median *tau*
0.10	0.09	0.11	0.10	0.08	0.10	0.09	0.12		0.09

GSPT/NCRA does not rely on inferential statistical analysis because BPRs are measured for every individual and GSPT's threshold-based approach. Developing the analytics to fit RDFs for all BPRs and HLTs was a substantial analytical challenge.

B. RESULTS

1. HLTp 1 Without Penalties

In this section, the procedures detailed in Chapter V, Section A.4 are documented, with exception of steps 3–5 as discussed. The reader is encouraged to focus on the unique fitting of the RDFs, agreement between forecasted and actual values, and the rich insights produced by the resource profile figures. Specifically, the variance observed across predominating BPRs at various levels of HLTp suggest potential capacity for understanding system resource issues.

a. Plot MBHSI Models and Examine for Errors (from Project I)

Figures 52 and 53 provide BPR vs. HLTp 1 Project I models without penalties enforced. Each plot identifies the performance vector and BPR (V41 = HLTp 1 without penalty). The x-axis is HLTp score and the y-axis is the original BPR capacity measurement. The majority of these models represent GSPT characteristics as described in Chapters III and IV. Note the decreasing density of data as the HLTp values increase along the x axis. These results suggest the HLT required high BPR capacities; that is, the task was difficult. In the majority of the models, there is a pattern of data increasing in value along both the x and y axes.

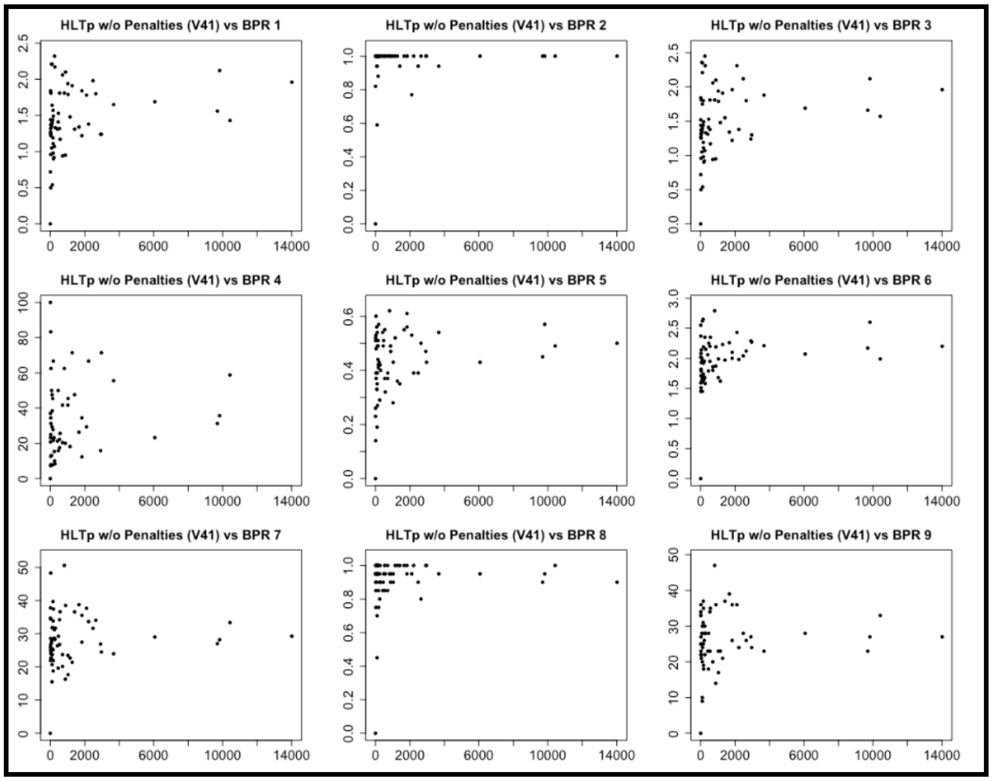


Figure 52. HLTp 1 without Penalties vs. BPRs 1–9

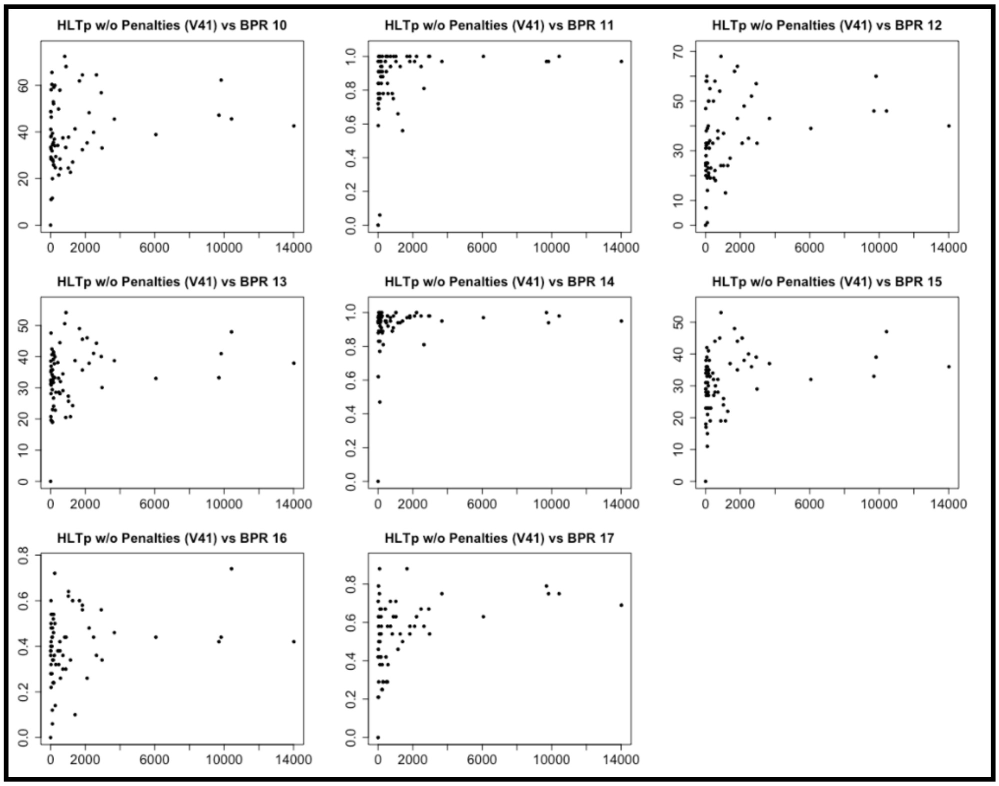


Figure 53. HLTp 1 without Penalties vs. BPRs 10–17

b. Construct RDFs Along Each BPR in the Model Set

Figures 54 and 55 provide all 17 BPR/HLTp 1 (without penalties enforced) models with RDFs. These RDFs use a set quantile (.07), fitting the RDF to encapsulate 93% of the data in each model. Desired NCRA characteristics were observed, with the possible exception of accuracy BPRs (BPRs 2, 8, and 14). In general, nonlinear climbing RDFs demonstrate these desired characteristics. The models for BPRs 2, 8, and 14 demonstrate the ceiling effect. If the ceiling effect (e.g., BPR 2 in Figure 54) results, the input (BPR) will likely result in under or over prediction of HLTp. The forecast may prematurely contact the RDF (under prediction) as the forecast moves across the model in search of the RDF or it may never contact it (over prediction). The practitioner should carefully consider inclusion of these models. This section included these models to highlight this critical issue in MBHSI performance forecasts. In the case where a stakeholder values HLT accuracy, which is the assumed case in this study, their inclusion is encouraged. Forecasts can be re-run to quantify errors regarding inclusion or inclusion. Review of the individual models and their identification as limiting BPRs (discussed in step 7) provides additional insight when determining inclusion in the model set. For example, if a BPR was not limiting for any participant, then removal may be justified. Also, a review of accuracy data as shown in step 6 when comparing predicted vs. actual values provides additional insight regarding model inclusion. For example, if a predominating limiting BPR is noted, yet the forecasted values vastly disagree with observed, then inclusion may be questioned. Finally, the models should, in general, demonstrate the characteristics detailed in Project I for inclusion in the model set. Recall that steps 3–5 results are documented in step 6.

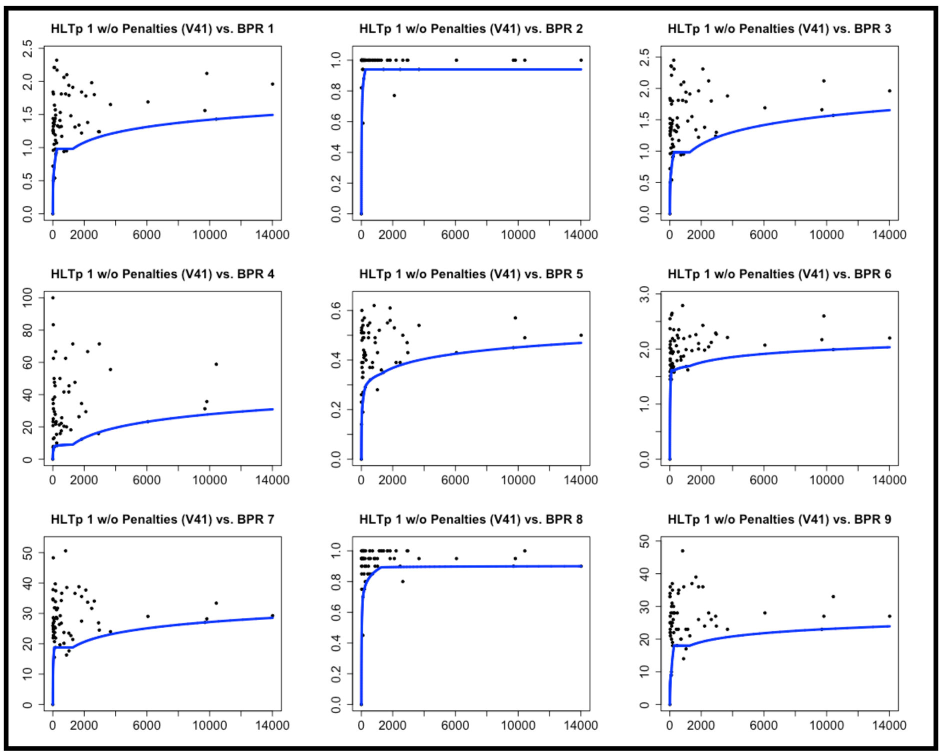


Figure 54. HLTp 1 without Penalties, RDFs (1–9)

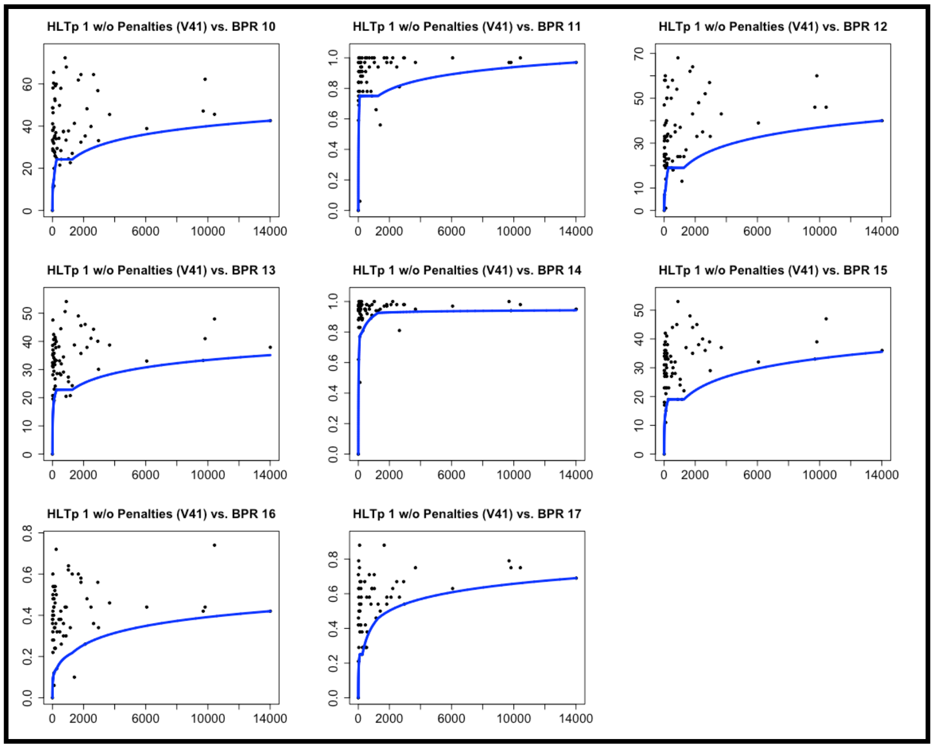


Figure 55. HLTp 1 without Penalties, RDFs (10–17)

c. *Appraise Difference Between Average Predicted HLTp 1 and Actual HLTp 1 Values*

This step documents the outcomes of steps 3–5. Figures 56 and 57 provide a brief review of the actual and predicted data. The boxplot in Figure 56 illustrates actual and predicted distributions. Figure 57 presents the actual vs. predicted data in a scatterplot after scaling ($x^{.3}$). Over and under prediction is identified by the red diagonal line. A Spearman's rank-order correlation evaluated the relationship between actual vs. predicted. There was strong positive (monotonic) correlation between actual and predicted HLTp, which was statistically significant ($r_s = .53$, $p < 0.001$).

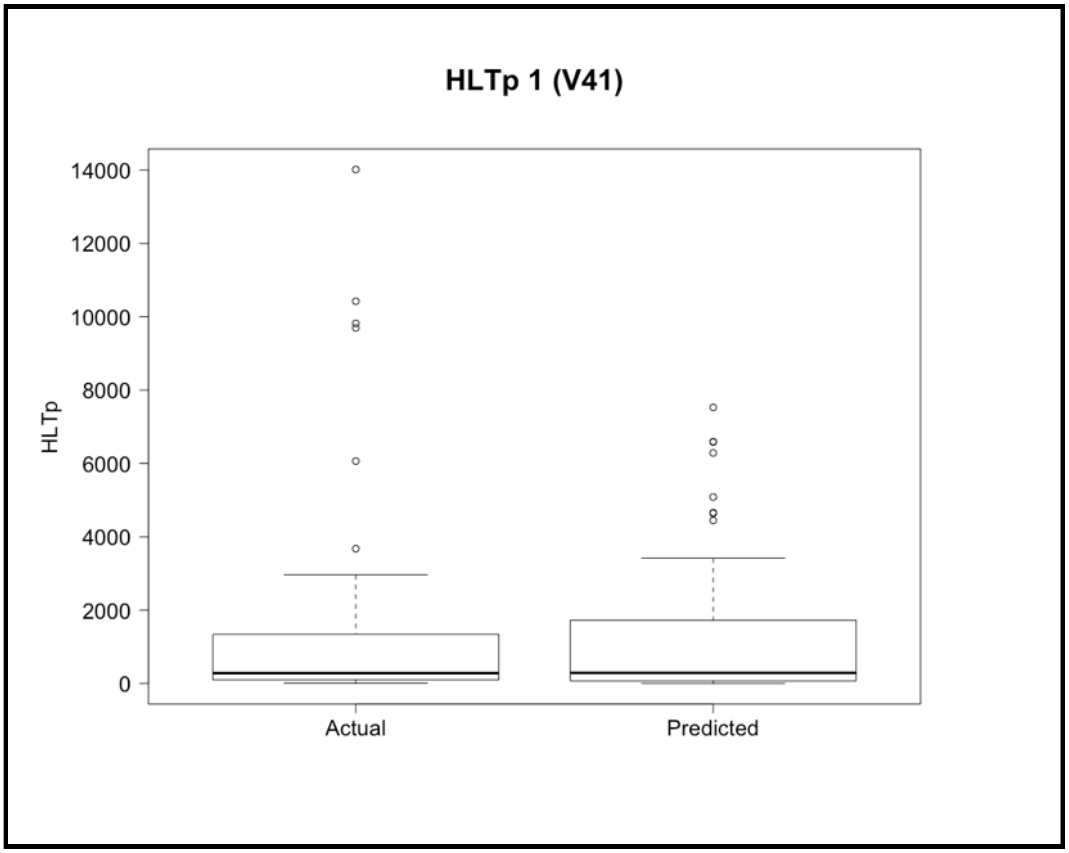


Figure 56. HLTp 1 (V41) Actual vs. Predicted Scores Boxplot

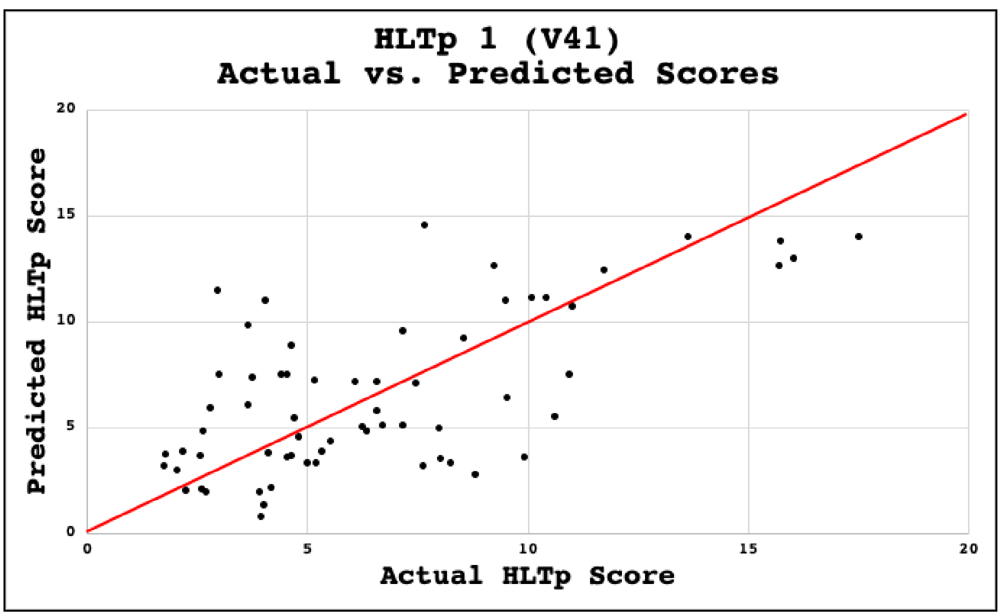


Figure 57. HLTp 1 (V41) Actual vs. Predicted Scatterplot Data After Scaling ($x^{.3}$)

Actual median HLTp, median predicted HLTp, error results, and accuracy are published in Table 9 for HLTp 1 (V41). Accuracy was calculated by dividing the absolute value of the difference between actual and predicted median performance by 282, the observed median HLTp. The accuracy rate (97%) also suggests the forecast models demonstrate agreement between predicted and actual performance.

Table 9. Median Observed and Predicted Scores

HLTp 1 (V41) - Median Scores		
Actual	**Predicted**	**Accuracy**
281.97	289.15	97%

Table 10 provides error rates for HLTp 1 using the median score of 282. Hits or misses are determined by the forecasted values being greater than the median. The false positive (Type I Error) and false negative (Type II Error) rates at this HLTp were both .34. The model accurately detected 42/64 participant outcomes (accuracy = .66).

Table 10. HLTp 1 Forecast Results for an HLTp of 282 (HLTp 1 Median)

HLTp 1 (V41) Score = 282	**Actual HLTp**	
Predicted	FALSE	TRUE
FALSE	21	11
TRUE	11	21
True/False Accuacy	**Type I Error Rate**	**Type II Error Rate**
66%	0.34	0.34

Table 11 demonstrates improved Type I (.09) error forecasts at an HLTp of 1,500. This score reflects the 75th percentile in observed HLTp scores. In total, the model accurately detected 54/64 participant outcomes (accuracy = .84). This evidence suggests that accuracy regarding errors may be influenced by HLTp and the RDF formulation, perhaps due to the skewness of the data or the absence of limiting BPRs in the study for some participants. If the HLTp set by the stakeholder results in unacceptable error rates, then adjustment to the RDF might improve error rates by trading off accuracy. It appears that higher HLTp values yield lower error rates. If higher HLTp values are set, one potential effect is the impact of reducing the number of personnel who would qualify for the HLTp, and that would impact the personnel and manpower (and probably the training and HFE) domains.

Table 11. Improved Type I and II Error Rates at HLTp 1,500

HLTp 1 (V41) Score = 1,500	**Actual HLTp**	
Predicted	FALSE	TRUE
FALSE	43	6
TRUE	4	11
True/False Accuacy	**Type I Error Rate**	**Type II Error Rate**
84%	0.09	0.35

Table 12 organizes predicted and actual HLTp scores in a matrix by categories. A Cramer's V was also run to determine strength of the diagonal in a square contingency table. The effect size (ES) was strong (ϕ\` = .48) (Cohen, 1988). A Cramer's V was also run using a 2x2 matrix to determine strength of the diagonal using the study median (*Mdn* = 282). The ES was moderate ($\phi$\` = .31) (Cohen, 1988).

Table 12. HLTp 1 Actual vs. Predicted Contingency Table

Cramer's V = .48	Actual HLTp Scores				
Predicted	< 200	200 – 499	500 – 999	1,000 – 4,999	5000 +
< 200	17	5	1	4	0
200 – 499	3	1	3	3	0
500 – 999	4	2	2	1	0
1000 – 4999	4	0	1	7	1
5000 +	0	0	1	0	4

Table 13 evaluates and documents actual HLT performance vs. absolute forecast error. Actual score categories reflect Table 12 but use smaller categories for absolute error in an effort to gauge how well the model performed across the HLTp continuum. The yellow and green data represent acceptable forecast performance. Red shaded regions reflect significant over prediction (N = 11/64). These findings suggest limiting BPRs for these particular individuals were not part of the study. This was investigated by forecasting the same study sample (N = 64) using various BPR densities.

Table 13. HLTp 1 Actual vs. Absolute Error

	Actual HLTp Score				
Absolute Error	< 200	200 – 499	500 – 999	1,000 – 4,999	5000 +
< 100	14				
101 – 200	4	4	3		
201 – 500	1	3	3	2	
501 – 1,000	4		1	5	1
1,000+	4		2	9	4

Table 14 details median HLTp forecasts using cumulative BPR groupings across all seventeen BPRs. The forecasts were executed using the same RDF criterion. As BPRs groups increase in size, forecast accuracy reliably recedes from over prediction toward the actual median. This finding suggests that as number of relevant BPRs increase, the model forecast performance also increases. This finding also suggests that some percentage of participants had a limiting BPR which were not included in the BPR test battery.

Table 14. BPR Density Effect on Forecast Performance

Cumulative BPR Density	HLTp 1 (V41) - Median Scores		
	Actual	Predicted	% Difference
BPR 1	281.97	8824.09	3029.1%
BPR 2	281.97	7463.3	2546.6%
BPR 3	281.97	5478.57	1842.8%
BPR 4	281.97	4123.6	1362.3%
BPR 5	281.97	2807.22	895.5%
BPR 6	281.97	1854.08	557.5%
BPR 7	281.97	1594.07	465.3%
BPR 8	281.97	797.98	183.0%
BPR 9	281.97	745.33	164.3%
BPR 10	281.97	594.1	110.7%
BPR 11	281.97	376.79	33.6%
BPR 12	281.97	376.79	33.6%
BPR 13	281.97	376.79	33.6%
BPR 14	281.97	289.15	2.5%
BPR 15	281.97	289.15	2.5%
BPR 16	281.97	289.15	2.5%
BPR 17	281.97	289.15	2.5%

d. Generate Pareto Data for BPRs across the HLT

Figure 58 illustrates each of the seventeen BPRs used in the study and the number of times each BPR limited participant HLTp. BPR 17 (Perceptual Integration Capacity-Snowy Pictures) predominated as the limiting BPR for HLTp 1 (N = 10). BPR accuracy tasks (2, 8, and 14) account for 13 participant-limiting BPRs. These BPRs were also those that demonstrated ceiling effect. In this model set, the accuracy models along with BPR 15 could be considered for removal when measuring forecast accuracy. But they deliberately remain in this model set to demonstrate these characteristics. A reason for inclusion could be a stakeholder priority of HLT accuracy, for example.

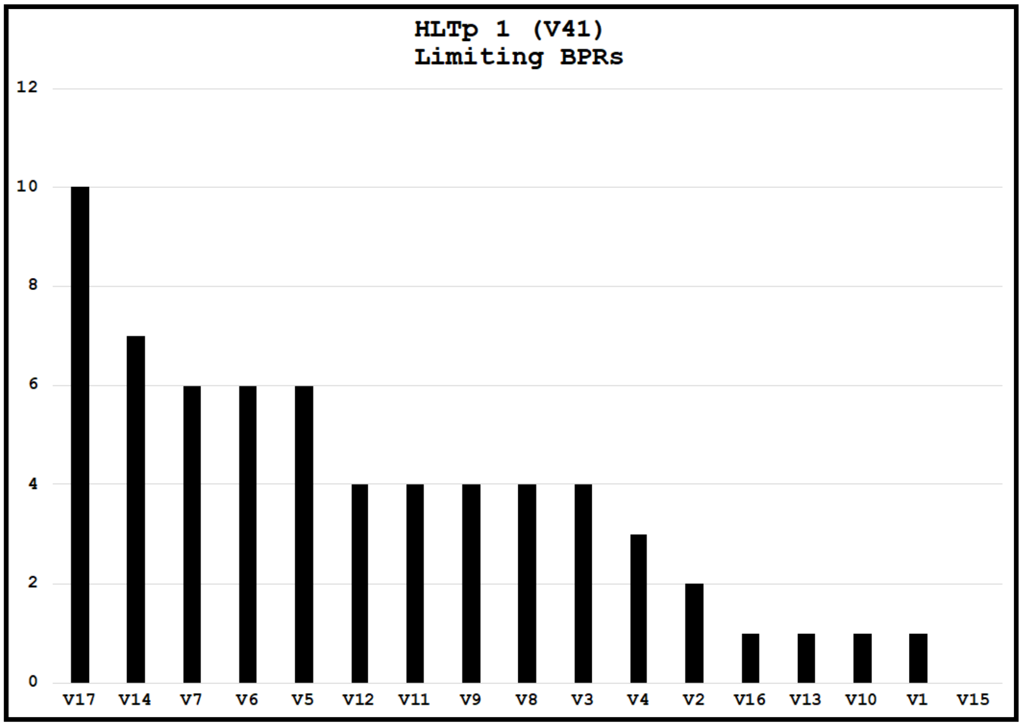


Figure 58. HLTp 1 Limiting BPRs

e. Construct Participant R_A Profiles

Figures 59 and 60 provide a sample of participant R_A profiles. These profiles represent the conditional expectation: E [HLTp | R_A]. Each profile lists forecasted performance by BPR. Recall in NCRA that the lowest forecasted performance value determines forecasted performance. The highest observed participant HLTp (14,015) serves as the maximum forecastable HLTp value. This limit bounds those forecasts that do not contact the RDF i.e., a high BPR capacity that when seeking the RDF never contacts it. Each participant demonstrated some ability across all seventeen BPRs in the study set. This suggests that humans in this study generally displayed the same set of abilities (Fleishman, 1992) but in different measures. Note the variation in BPR order on the y-axis. For example, participant 3 BPR capacities forecast an HLTp of just 69 because of the participant's R_A for V2, but the limiting resource for participants 1 and 4 were V5 and V14, respectively.

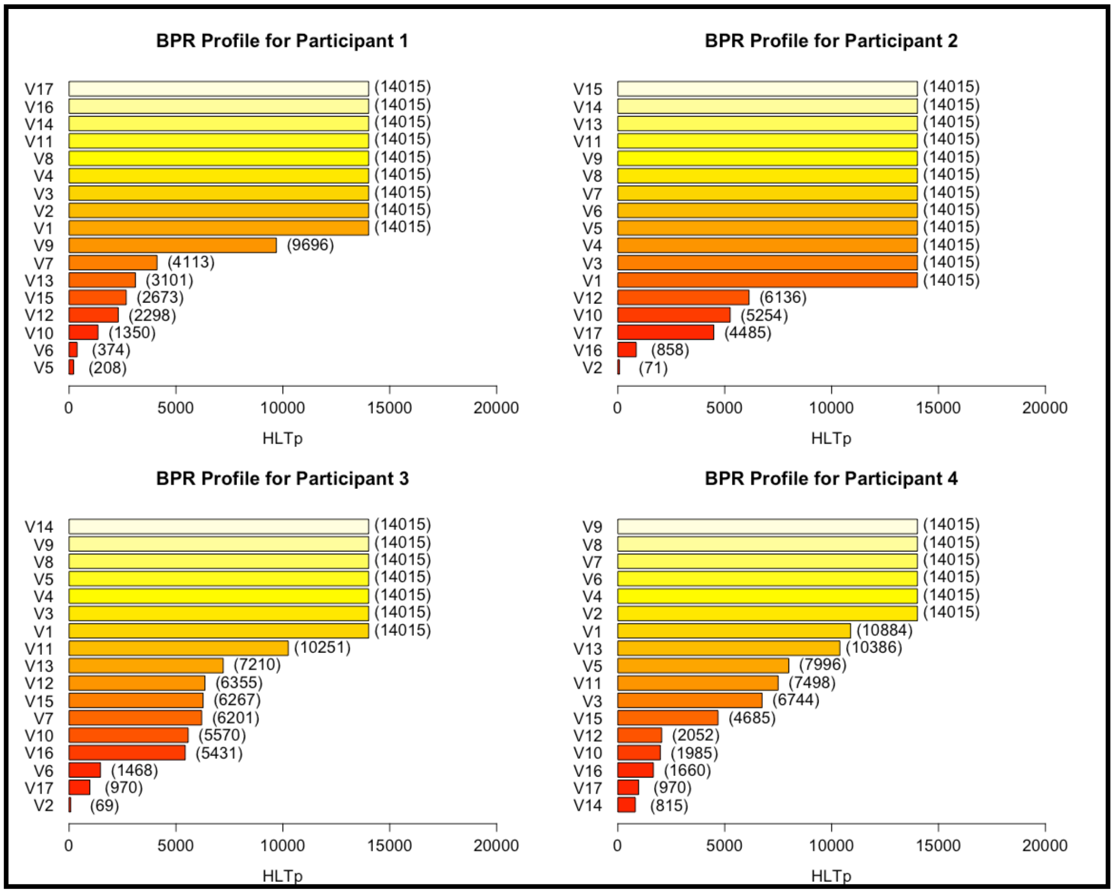


Figure 59. Sample R_A Profiles

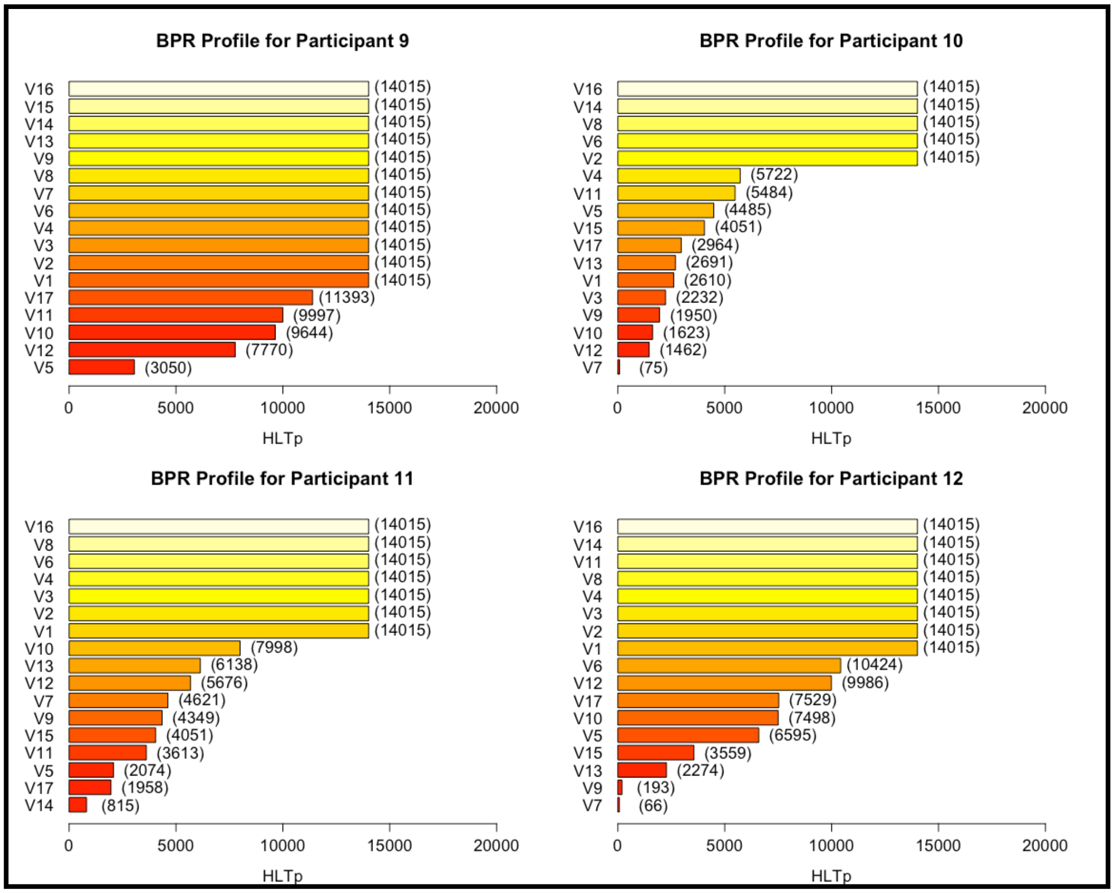


Figure 60. Sample R_A Profiles

f. Construct R_D Profile for a Set of HLTp Values

Figure 61 provides a sample of R_D profiles for performance thresholds of 150, 282, 1500, and 5000. These profiles represent the conditional expectation: E [R_A | HLTp]. Each profile lists forecasted threshold-levels of BPRs by HLTp. Each HLTp value documents the minimum capacity expected across each of the seventeen BPRs for a given level of HLTp. Specific BPR values (as measured in the study) and respective percentiles (x-axis) standardize the various BPR measurements. Actual minimum BPR score and percentile establish the profile requirements. The order of the BPRs on the y-axis changes from one HLTp to the next indicating that as the performance level changes so do the requisite BPRs scores and percentiles. Considering this, it would be challenging to forecast the system design trade space without these data. Another key insight to these profiles suggests that dominance in one or a few BPRs is not necessary; rather, minimal capacity across each BPR is required for most HLTp levels.

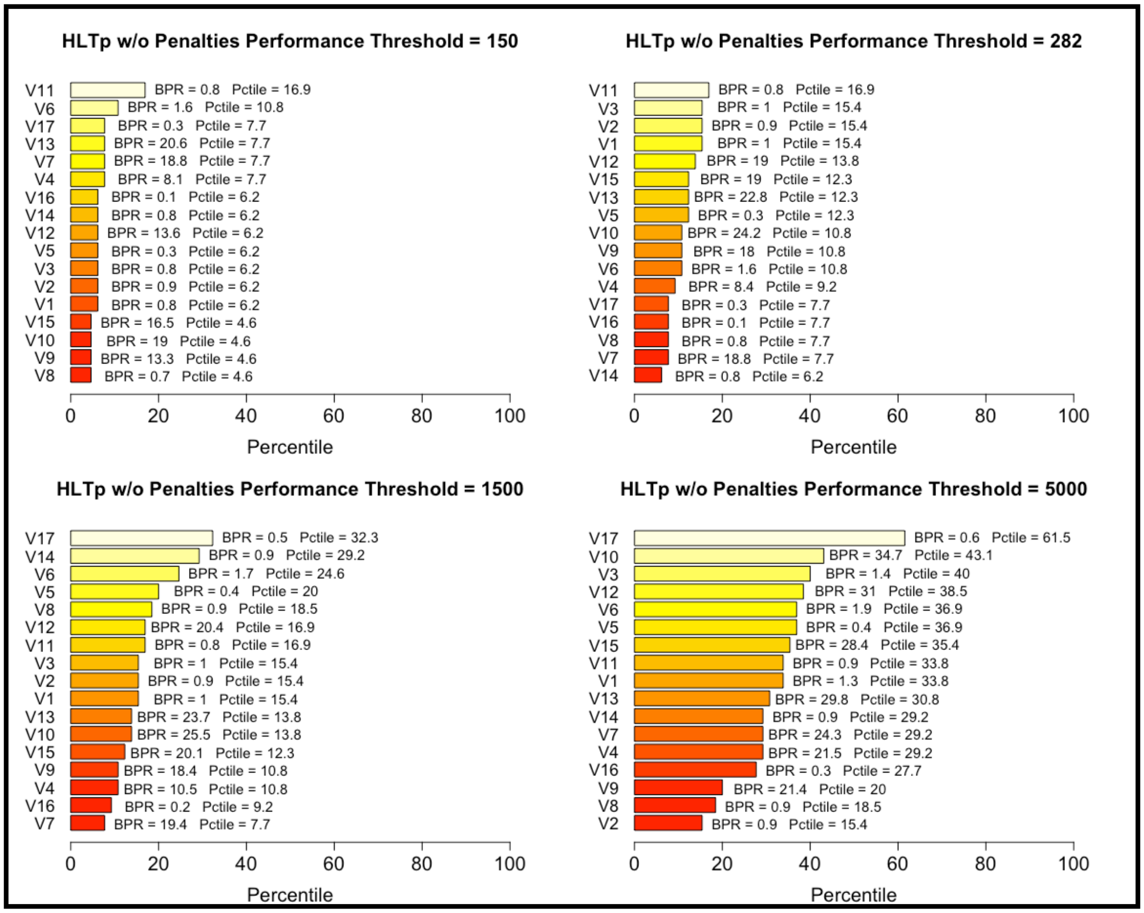


Figure 61. Sample R_D Profiles at Various HLTp Scores

In summary, this systemic, iterative, and systematic approach demonstrated in this section (steps 1–9) is a representative example of MBHSI's value proposition to MBSE. Specifically, this example documents how MBHSI models are built, forecasts are executed, accuracy is measured, and HSI resources (inputs) and performance (outputs) are quantified. The quantification of R_A and R_D, support establishment of initial constraint values for optimization program formulation discussed in Chapter VII. Additionally, the data outputs from this process produce insights regarding limiting BPRs, predominance of BPR limitation within the sample set, and resource limitations at the individual level. Project III uses these data as baselines to measure effects of training and automation on HLTp. Chapter VII then addresses how these values translate to mathematical program formulation in terms of optimization. Next, similar evaluations of the sub-HLTs are presented individually to demonstrate this capacity at different levels of the HLT hierarchy.

2. Airspeed Sub-HLTp 1

This section repeats the Chapter V, Section A.9 procedures using Airspeed sub-HLTp 1. This demonstration examines the capacity for MBHSI to execute the process at different levels of the task hierarchy. Recall that maintaining Airspeed represented one of the three HLT subtasks. The reader is encouraged to continue to focus on the unique fitting of the RDFs, agreement between forecasted and actual values, and the rich insights produced by the resource profile figures. Specifically, the continued variance observed across predominating BPRs at various levels of HLTp suggest improved capacity for understanding system resource issues.

a. Plot MBHSI Models and Examine for Errors (from Project I)

Project I contains the Airspeed models. The models in step 2 are fitted with RDFs.

b. Construct RDFs along Each BPR in the Model Set

Figures 62 and 63 provide all seventeen BPR/Airspeed models with RDFs. The RDFs encapsulate 92% of the data in each model (quantile = .08). The Airspeed models demonstrate agreement with NCRA concepts. As expected, BPRs 2, 8, 11, and 14 demonstrate ceiling effects just as they did in HLTp 1. The x-axis in each model represents the sub-HLTp score and the y-axis represents original BPR capacity data (not scaled) unique to each BPR measurement.

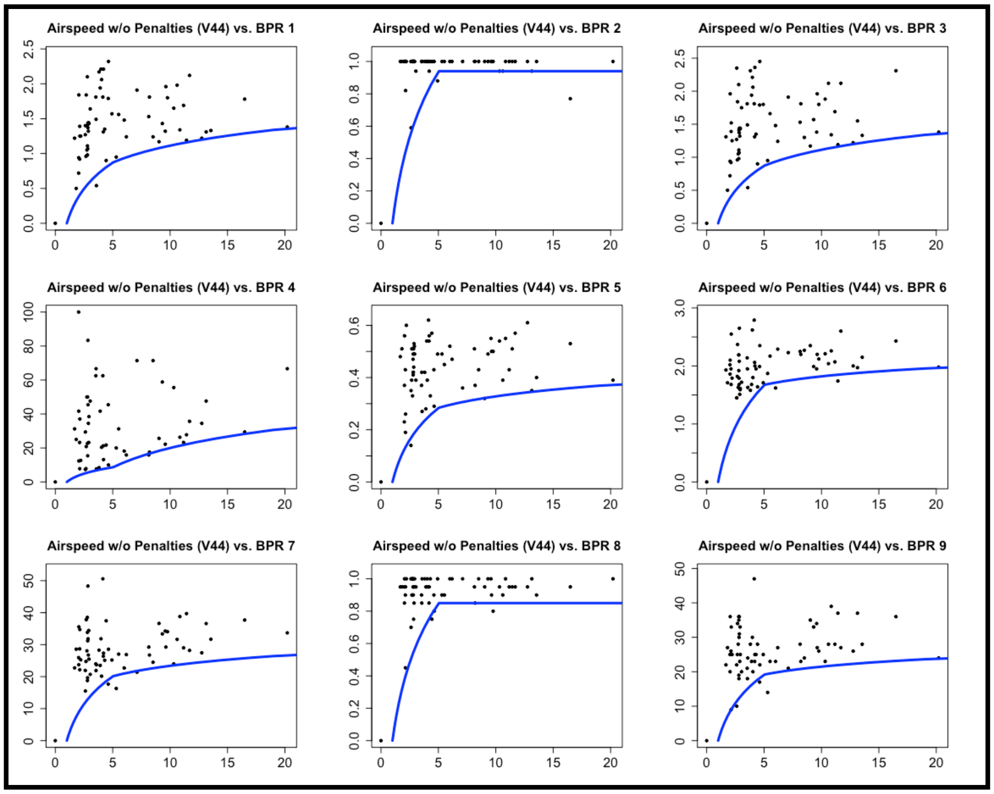


Figure 62. Airspeed 1 without Penalties RDFs (1–9)

Figure 63 BPR models (10, 15, 16, and 17) demonstrate difficulty encapsulating higher HLTp data. The relatively few data points towards the right side of the models have reduced leverage on the RDF. The novel MBHSI R code logic can adjust for this phenomenon as shown in Figure 64. However, the limited data reduce confidence in accuracy because so few data exist at the extreme edge of this data set. Adjustment may result in a more sporadic RDF near the lower end of the HLTp data. This suggests over-fitting of the RDF. In the event the stakeholder desires a higher HLTp, this approach may provide some insight. This emphasizes the importance of determining the most appropriate HLTp value up front. Early knowledge of the target HLTp value can focus the researcher on obtaining more data points near the higher HLTp value. The RDFs in Figure 64 also illustrate the issue of model resolution vs. power analysis discussed in detail in Chapter II.

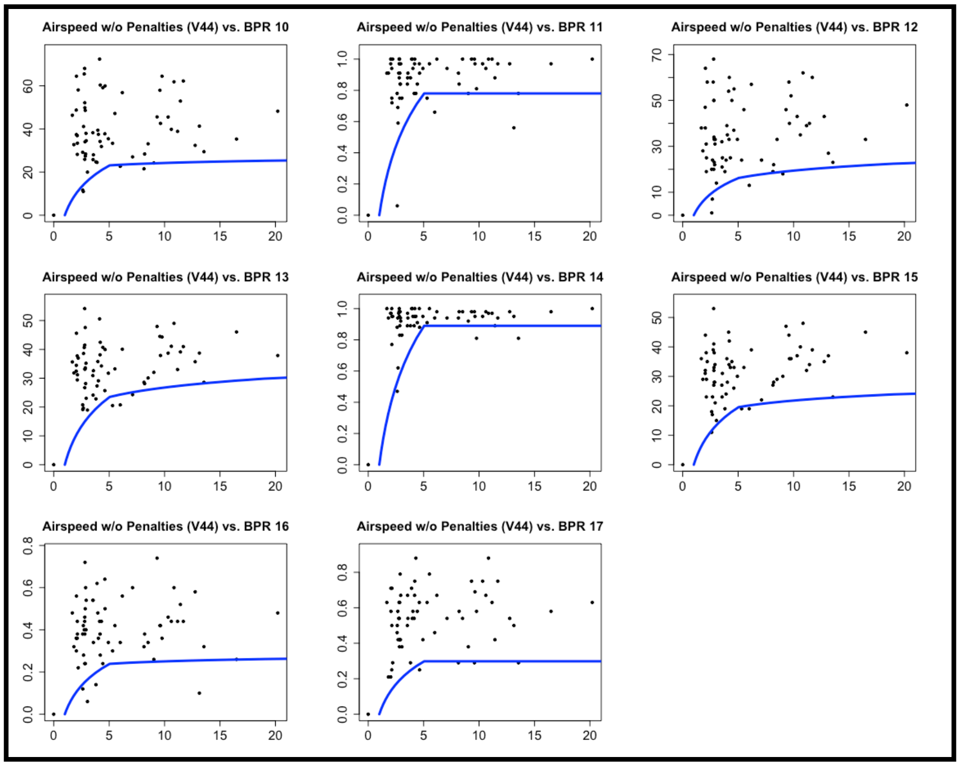


Figure 63. Airspeed 1 without Penalties RDFs (10-17)

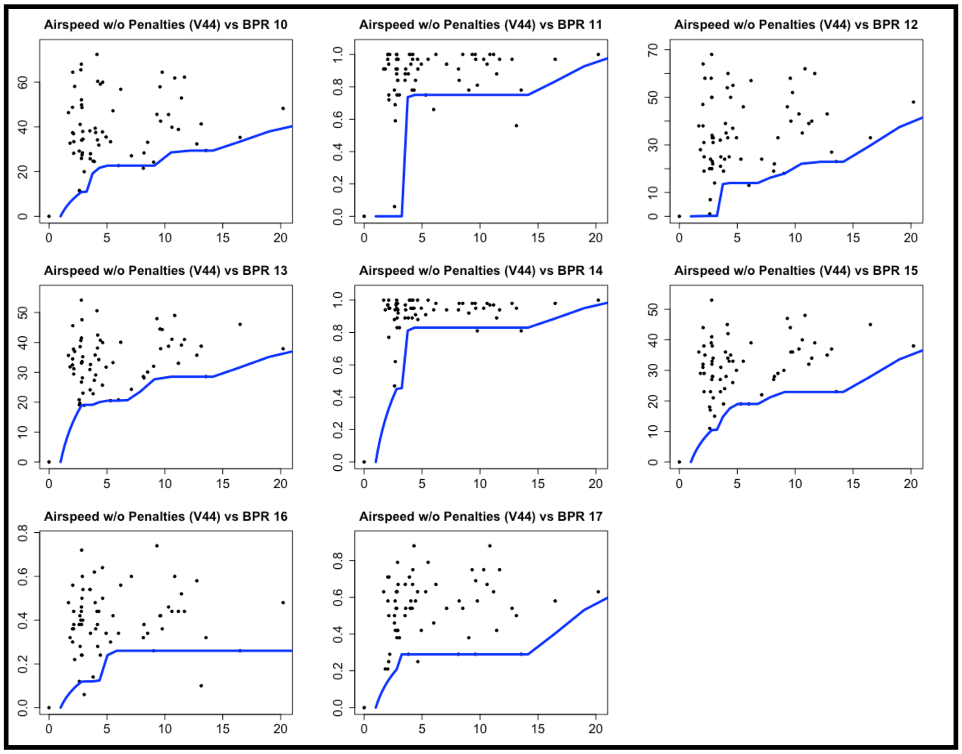


Figure 64. Airspeed Over-Fit RDF (10–17)

3. Appraise Difference between Average Predicted HLTp and Actual HLTp Values

Step 6 documents the outcomes of steps 3–5. Figures 65 and 66 provide a brief review of the actual and predicted data. The boxplot in Figure 65 illustrates actual and predicted distributions. Figure 66 presents the actual vs. predicted data in a scatterplot. Over and under prediction is identified by the red diagonal line. A Spearman's rank-order correlation evaluated the relationship between actual vs. predicted. There was moderate positive (monotonic) correlation between actual and predicted HLTp, which was statistically significant ($r_s = .37$, $p = 0.003$).

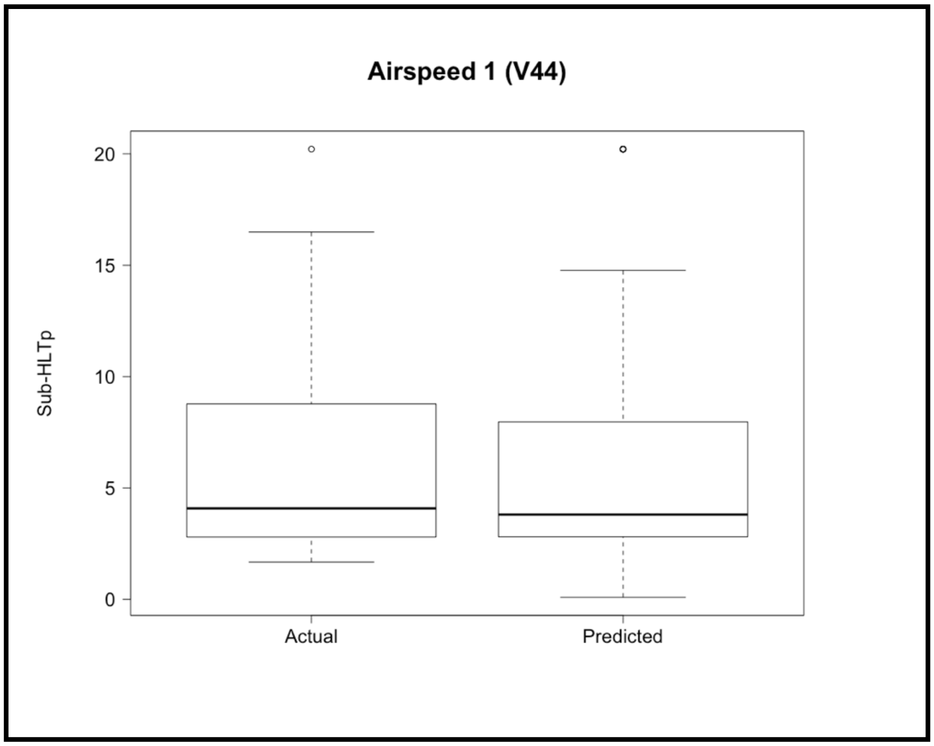


Figure 65. Airspeed (V44) Actual vs. Predicted Scores Boxplots

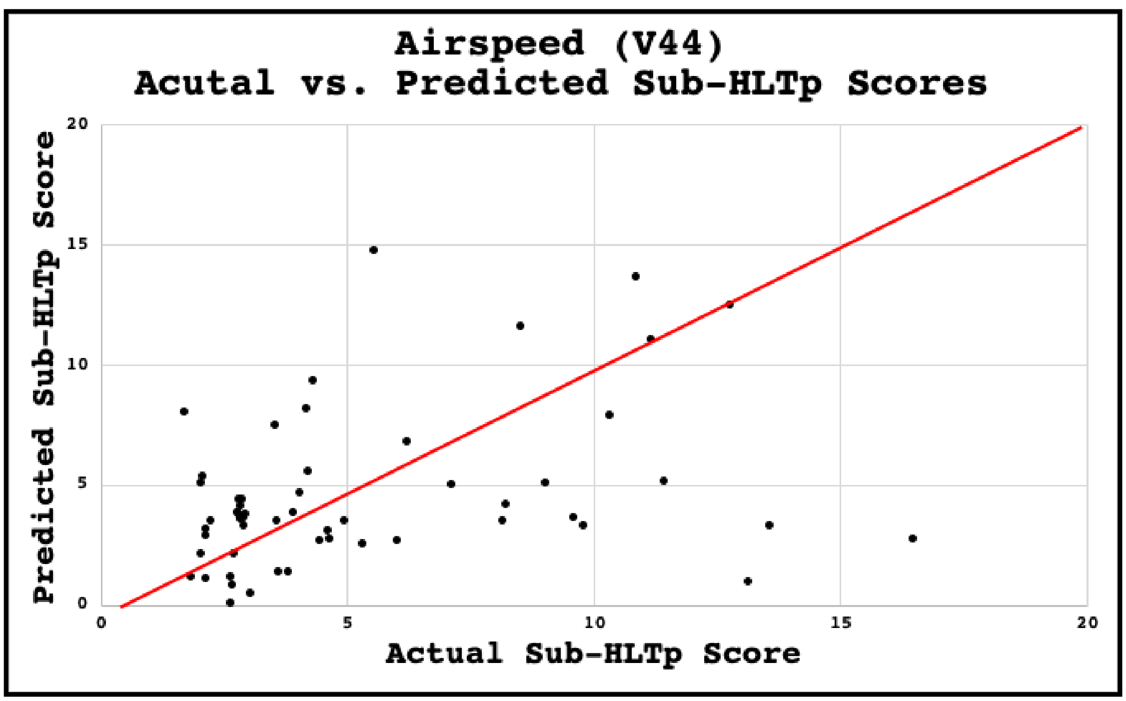


Figure 66. HLTp 1 (V41) Actual vs. Predicted Scores Scatterplot (No Scaling)

Table 15 publishes actual median Sub-HLTp, median predicted HLTp, error results, and accuracy using Airspeed observed score (*Mdn* = 4.08). The performance forecast accuracy rate (93%) suggests that the forecast models demonstrate agreement between predicted and actual performance.

Table 15. Airspeed Forecast Accuracy

Airspeed 1 (V44) - Median Scores		
Actual	**Predicted**	**Accuracy**
4.08	3.8	93%

Table 16 provides the model's Type I error rates using the median Airspeed score. Adjustment of the RDF may improve error avoidance but forecast accuracy may degrade. The priorities of the stakeholders determine the proper RDF fit for the system.

Table 16. Airspeed 1 Forecast Results Using the Median

Airspeed 1 (V44) Score = 4	Actual Sub-HLTp 1 Score	
Predicted	FALSE	TRUE
FALSE	22	9
TRUE	12	21
True/False Accuacy	**Type I Error Rate**	**Type II Error Rate**
67%	0.35	0.30

Table 17 demonstrates improved Type I (.22) and degraded Type II (.44) error forecasts at a score of 7. In total, the model accurately detected 46/64 participant outcomes (accuracy = .72). This evidence further supports the assertion made in the prior section (HLTp) that accuracy regarding errors may be influenced by HLTp level and the RDF formulation, perhaps due to the skewness of the data or the absence of limiting BPRs in the study for some participants. This model suggests that relatively more Type II errors were noted at the higher Airspeed score of 7. These results are due to under-prediction of performance at the higher HLTp levels of the sub-HLT. The challenging nature of this sub-HLT resulted in fewer high scores thus, the RDF accuracy may be limited with such few data occupying this portion of the model.

Table 17. Type I and Airspeed I Error Rates, Score = 7

Airspeed 1 (V44) Score = 7	**Actual Sub-HLTp 1 Score**	
Predicted	FALSE	TRUE
FALSE	36	8
TRUE	10	10
True/False Accuacy	**Type I Error Rate**	**Type II Error Rate**
72%	0.22	0.44

Table 18 organizes predicted and actual Airspeed 1 scores in a matrix by categories. A Cramer's V was also run to determine strength of the diagonal in a square contingency table. The effect size (ES) was strong (ϕ` = .3) (Cohen, 1988). A Cramer's V was also run using a 2x2 matrix to determine strength of the diagonal using the study median (*Mdn* = 4). The ES was moderate ($\phi$` = .35) (Cohen, 1988).

Table 18. Airspeed 1 Actual vs. Predicted Contingency Table

Cramer's V = .3	**Actual Sub-HLTp 1 Scores**				
Predicted	< 2.9	3.0-5.9	6.0-8.9	9.0-12.9	13.0+
< 2.9	8	6	1	0	2
3.0-5.9	12	8	3	4	1
6.0-8.9	1	2	1	1	0
9.0-12.9	0	1	1	2	1
13.0+	1	3	0	5	0

Table 19 evaluates and documents actual sub-HLT performance vs. absolute forecast error. Actual score categories reflect Table 18 but use smaller categories for absolute error in an effort to gauge how well the model performed across the sub-HLTp continuum. The yellow and green data represent acceptable and desired forecast performance. The majority of participant forecasts fell within the acceptable zones (N = 58/64). Recall the range of actual scores (0-20.2) observed out of a possible range of 0 - 100. Red shaded regions reflect significant over prediction (N = 6/64). These findings suggest limiting BPRs for these particular individuals were not part of the study as demonstrated in the last section.

Table 19. Airspeed Actual vs. Absolute Error

	Actual Sub-HLTp 1 Scores				
Absolute Error	< 2.9	3.0-5.9	6.0-8.9	9.0-12.9	13.0+
< 2	19	8	1	1	1
2.1 - 4	3	6	4	3	
4.1 - 6			1	1	
5.1 - 8	1	1		2	
8.1 +	2	2		4	4

a. Generate Pareto Data for BPRs across the HLT

Figure 67 illustrates each of the 17 BPRs used in the study and how many times each BPR limited participant Airspeed performance. BPR 1 (Multi-Limb Coordination Speed) predominated as the limiting BPR for Airspeed performance (N = 12). Given the nature of controlling the aircraft and throttle using separate limbs suggests this predominating BPR may be expected. BPRs V15, V11, and V3 did not predominate as a limiting BPR for any of the 64 participants for this sub-HLT using this particular RDF fit. The R_A profiles extrapolate these data to the individual level, suggesting potential utility in system design, training, and personnel selection.

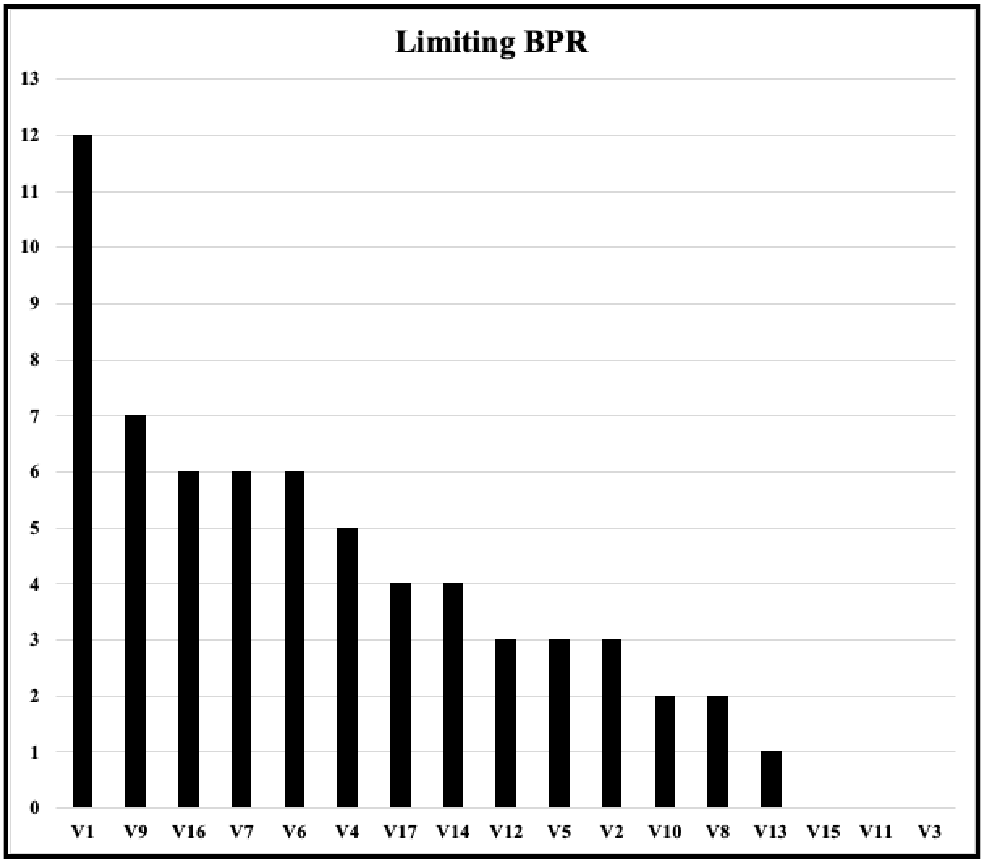


Figure 67. Airspeed Limiting BPRs

b. Construct Participant R_A Profiles

Figures 68 and 69 provide a sample of participant R_A profiles including forecasted performance for each BPR. These profiles represent the conditional expectation: E [HLTp | R_A]. Recall that in NCRA, the lowest forecasted performance value determines forecasted performance. For example, participant 1 BPR capacities forecasted an Airspeed performance of 4. The highest observed participant performance (20.2) served as the maximum forecastable value. BPR predominance by participant demonstrates additional insights at the sub-task level of the study.

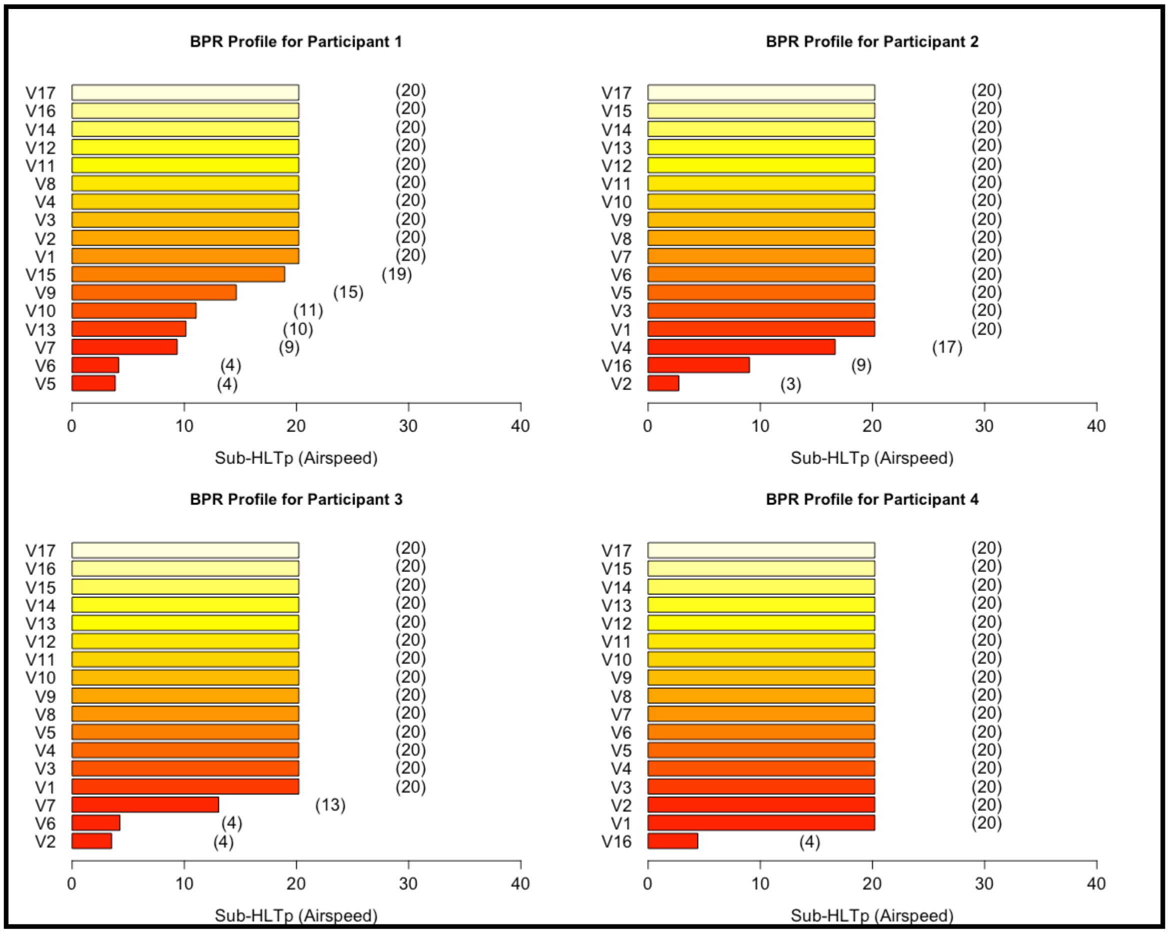


Figure 68. Sample Airspeed R_A Participant Profiles (1–4)

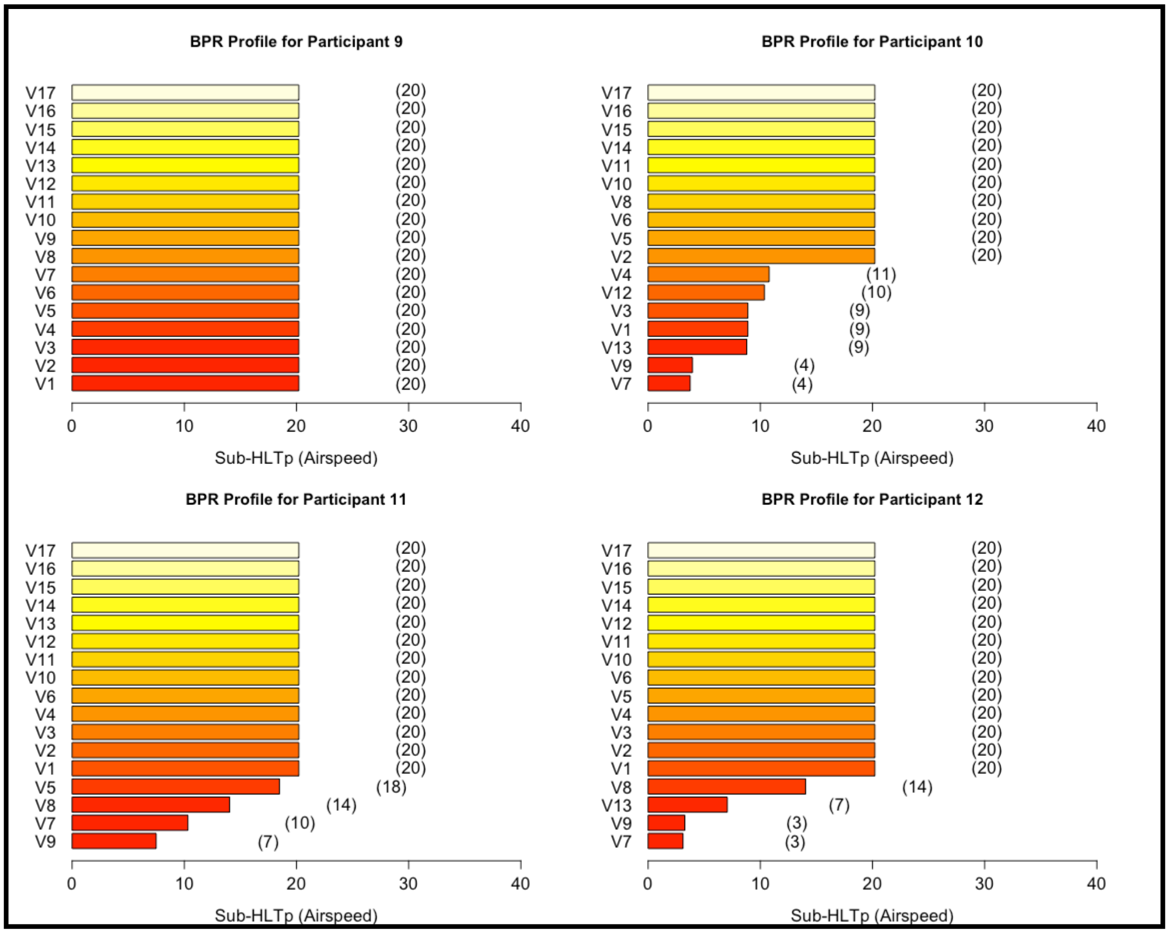


Figure 69. Sample Airspeed RA Participant Profiles (9–12)

c. *Construct R_D Profile for a Set of HLTp Values*

Figure 70 provides a sample of Airspeed R_D profiles, to include the median (*Mdn* = 4). These profiles represent the conditional expectation: E [R_A | HLTp]. Each profile lists forecasted threshold levels of BPRs by HLTp and each figure documents the minimum capacity expected across each of the 17 BPRs for a given level of Airspeed performance. Again, different BPRs predominate at various levels of performance. This evidence at different levels of the system hierarchy may prove to be of high value in supporting SE.

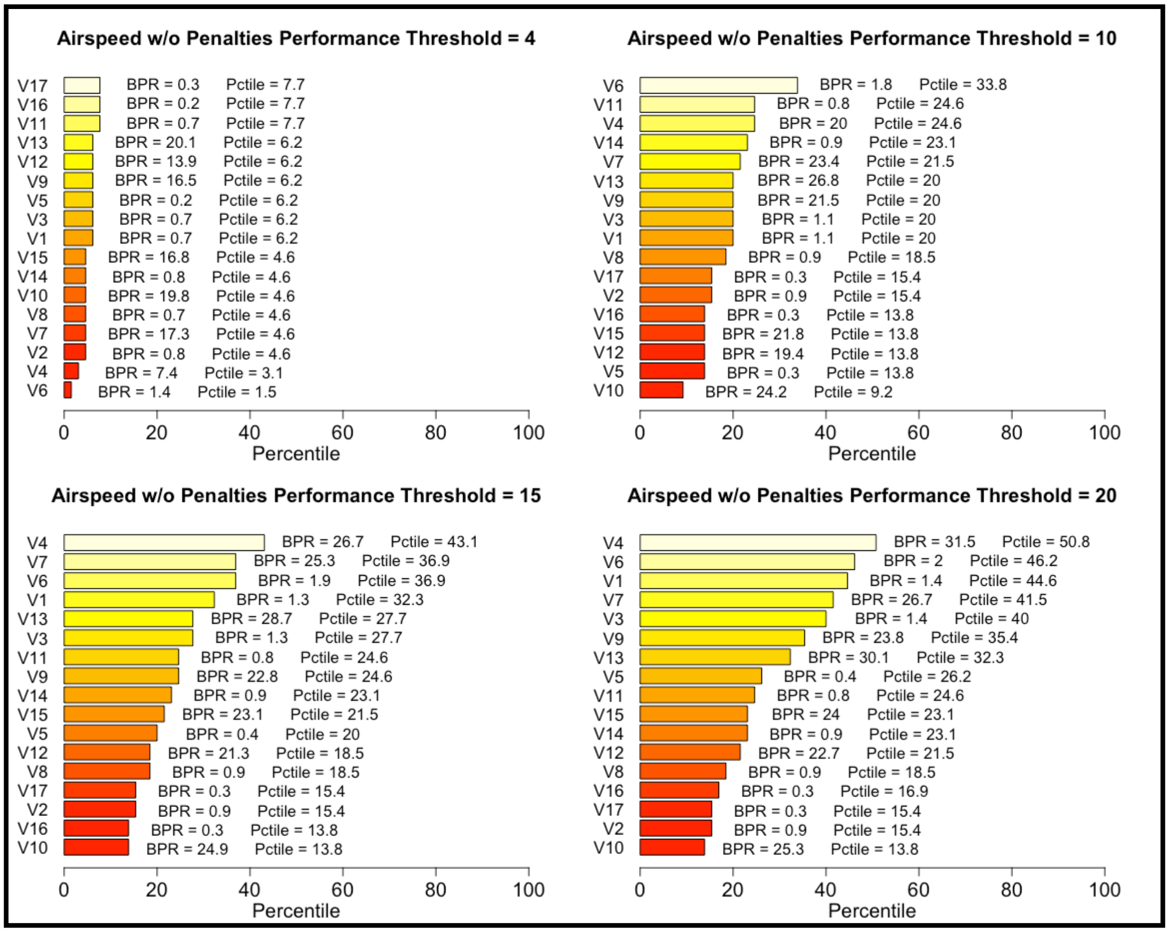


Figure 70. Airspeed R_D Profiles

In summary, the systemic, iterative and systematic approach demonstrated again in this section (steps 1–9) at the sub-HLT level is another representative example of MBHSI's value proposition to MBSE. Specifically, this example documents how MBHSI sub-HLT models are built, forecasts are executed, accuracy is measured, and HSI resources (inputs) and performance (outputs) are quantified. The quantification of sub-HLT R_A and R_D, further support establishment of initial constraint values for optimization program formulations discussed in Chapter VII. Additionally, the sub-HLT data outputs from this process produce insights regarding limiting BPRs, predominance of BPR limitation within the sample set, and resource limitations at the individual level. Next, consistent evaluations of the other two sub-HLTs are presented to demonstrate this capacity laterally across levels of the HLT hierarchy.

4. Course Sub-HLTp 1 (CDI-H)

Course sub-HLTp 1 represents one of the three HLT subtasks. This section also closely follows Chapter V, Section 3.A procedures. This demonstration further examines the capacity for MBHSI to execute the process at different levels of the task hierarchy. Additionally, this subtask demonstrates the capacity of MBHSI to execute the process across lateral subtasks of an HLT.

a. Plot MBHSI Models and Examine for Errors (from Project I)

Project I contains Course models. The models in step 2 include these models with fitted RDFs.

b. Construct RDFs along Each BPR in the Model Set

Figures 71 and 72 provide all 17 BPR/Course models with RDFs. The RDFs encapsulate 93% of the data in each model (quantile = .07). These models demonstrated better agreement when slightly over-fitting the RDFs as discussed in the prior section (Airspeed). When under-fit, poor fitting is noted on the right-side the RDFs. This characteristic may be explained by this subtask demanding fewer resources to a point. Specifically, many models (BPRs 10–17) demonstrate a plateau to nearly a score of 40, then an increase in resources can be detected. Tighter RDF fitting recognized this slope change in R_D. If not accomplished, then greater over-prediction of performance can be expected. Again, fitting should be defined by stakeholder priorities. In this case, the RDFs demonstrate a priority of forecast accuracy. The Course models continue to demonstrate excellent agreement with NCRA concepts.

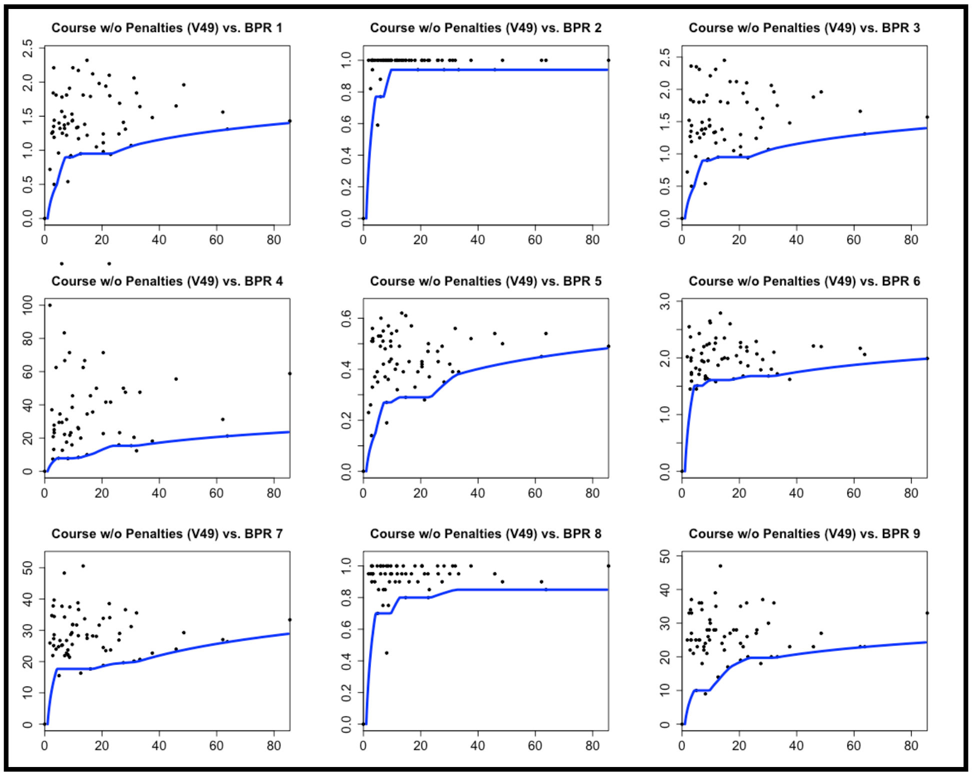


Figure 71. Course without Penalties, RDFs (1–9)

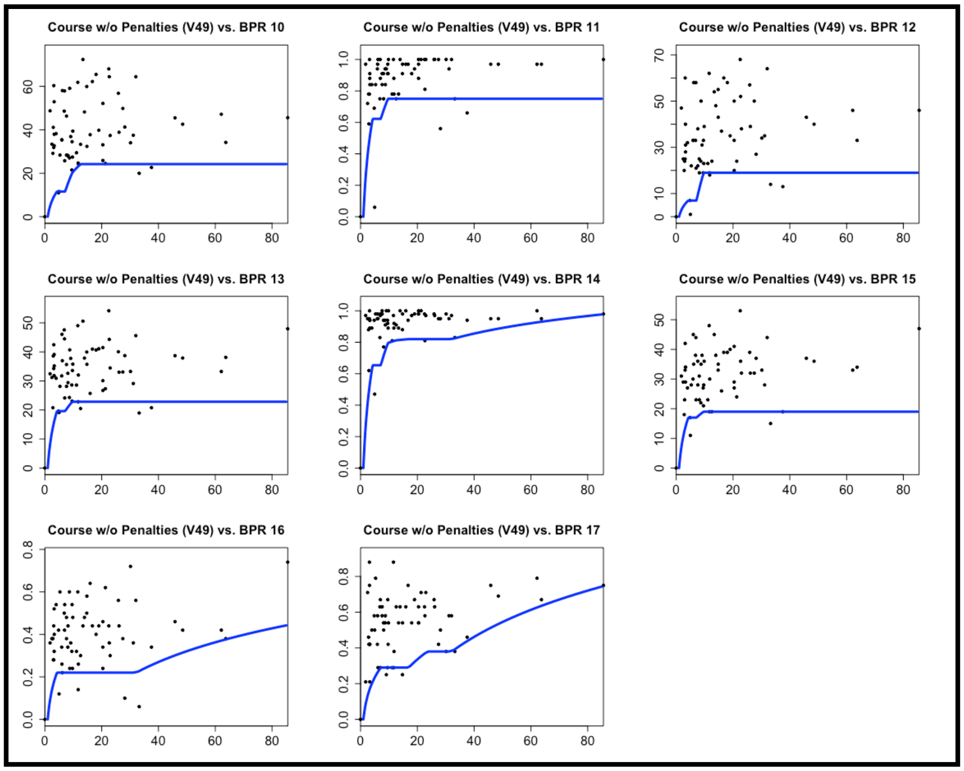


Figure 72. Course without Penalties, RDFs (10–17)

c. Appraise Difference between Average Predicted HLTp and Actual HLTp Values

Step 6 documents the outcomes of steps 3–5. Figures 73 and 74 provide a brief review of the actual and predicted data. The boxplot in Figure 73 illustrates actual and predicted distributions. Figure 74 presents the actual vs. predicted data in a scatterplot. Over and under prediction is identified by the red diagonal line. A Spearman's rank-order correlation evaluated the relationship between actual vs. predicted. There was moderate positive (monotonic) correlation between actual and predicted HLTp, which was statistically significant ($r_s = .36$, $p = 0.003$).

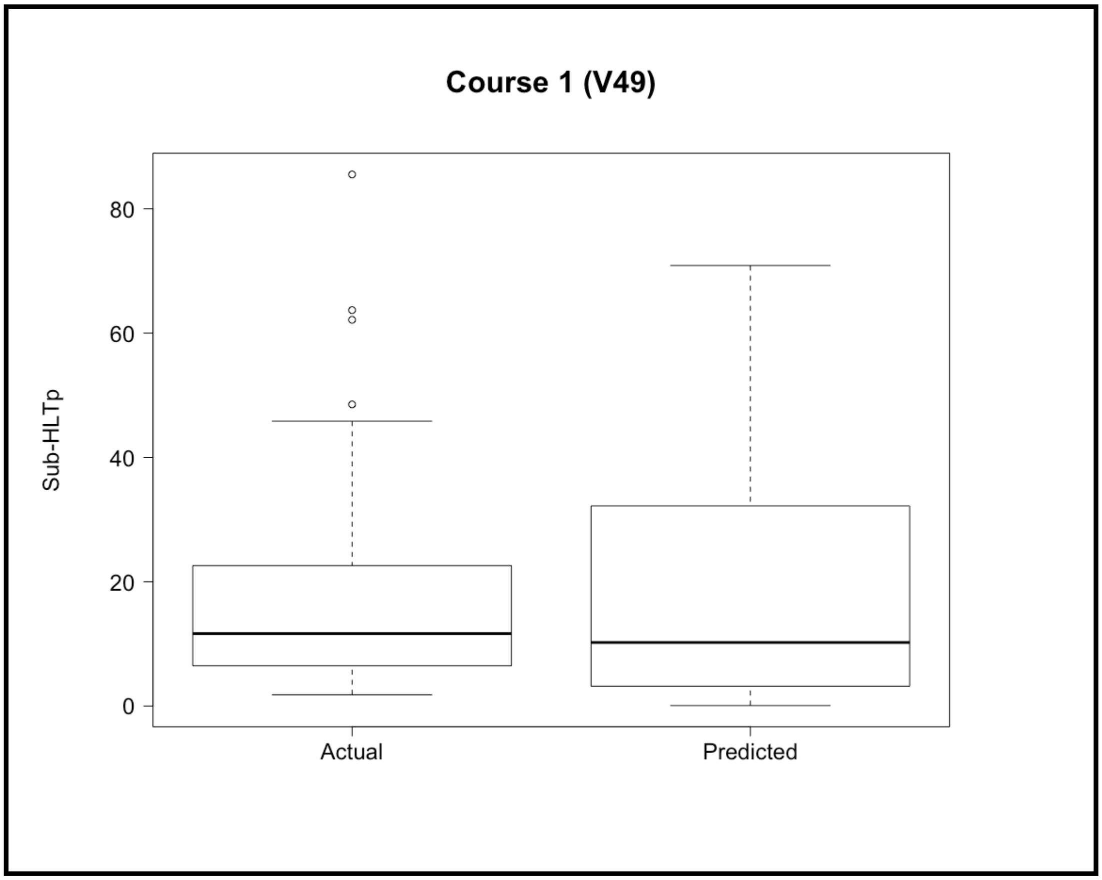


Figure 73. Course Actual vs. Predicted Scores Boxplots (No Scaling)

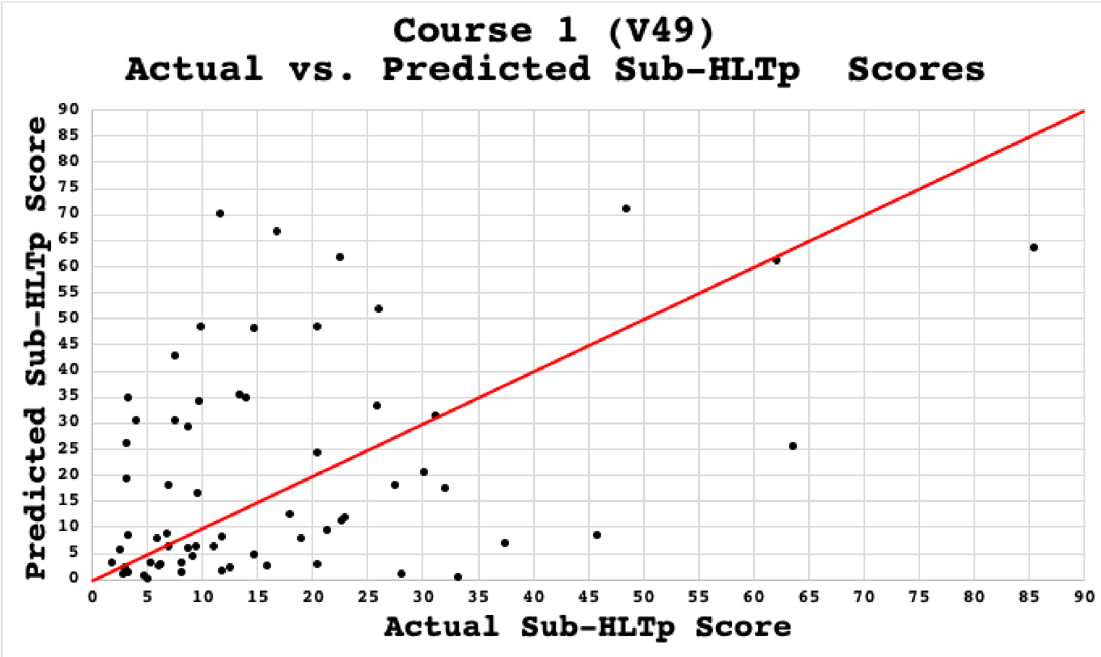


Figure 74. Course Actual vs. Predicted Scores Scatterplot (No Scaling)

Table 20 publishes actual median Sub-HLTp, median predicted HLTp, error results, and accuracy using Course observed score (*Mdn* = 11.65). The performance forecast accuracy rate (88%) suggests that the forecast models demonstrate agreement between predicted and actual performance.

Table 20. Course Forecast Accuracy

Course 1 (V49) - Median Scores		
Actual	Predicted	Accuracy
11.65	10.23	88%

Table 21 provides the model's error rates using the median Course score. Adjustment of the RDF may improve error avoidance but forecast accuracy may degrade. The priorities of the stakeholders determine the proper RDF fit for the system.

Table 21. Course 1 Forecast Results Using the Median

Course 1 (V49) Score = 11.65	Actual Sub-HLTp Score	
Predicted	FALSE	TRUE
FALSE	20	12
TRUE	13	19
True/False Accuacy	**Type I Error Rate**	**Type II Error Rate**
61%	0.39	0.39

Table 22 demonstrates improved Type I (.17) and degraded Type II error forecasts (.65) at a score of 25. In total, the model accurately detected 42/64 participant outcomes (accuracy = .66). This evidence further supports the assertion made in the prior section (HLTp) that accuracy regarding errors may be influenced by HLTp level and the RDF formulation. The Type II error results are due to under-prediction of performance at the higher levels of the sub-HLT. The challenging nature of this sub-HLT resulted in fewer high scores thus, the RDF accuracy may be limited with such few data occupying the higher performance areas of the model.

Table 22. Course Type I and II Error Rates with a Score = 25

Course 1 (V49) Score = 25	Actual Sub HLTp Score	
Predicted	FALSE	TRUE
FALSE	35	15
TRUE	7	7
True/False Accuacy	**Type I Error Rate**	**Type II Error Rate**
66%	0.17	0.65

Table 23 organizes predicted and actual Course 1 scores in a matrix by categories. A Cramer's V was also run to determine strength of the diagonal in a square contingency table. The effect size (ES) was strong (ϕ` = .3) (Cohen, 1988). A Cramer's V was also run using a 2x2 matrix to determine strength of the diagonal using the study median (*Mdn* = 11.65). The ES was moderate (ϕ` = .22) (Cohen, 1988).

Table 23. Course 1 Actual vs. Predicted Contingency Table

Cramer's V = .3	**Actual Sub-HLTp Score**				
Predicted	< 5	5.0-14.9	15-24.9	25-39.9	40+
< 5	6	9	2	2	0
5.0-14.9	2	7	5	1	1
15-24.9	1	2	1	3	0
25-39.9	3	5	0	2	1
40+	0	4	3	1	3

Table 24 evaluates and documents actual sub-HLT performance vs. absolute forecast error. Actual score categories reflect Table 23, but use smaller categories for absolute error in an effort to gauge how well the model performed across the sub-HLTp continuum. The yellow and green data represent acceptable and desired forecast performance. The majority of participant forecasts fell within the acceptable zones (N = 49/64). Recall the range of actual scores (0-86) observed out of a possible range of 0 -100. Red shaded regions reflect significant over prediction (N = 15/64). These findings suggest limiting BPRs for these particular individuals were not part of the study as demonstrated in the last section.

Table 24. Course Actual vs. Absolute Error

	Actual Sub-HLTp Score				
Absolute Error	< 5	5.0-14.9	15-24.9	25-39.9	40+
< 5	8	10	1	1	1
5.0 - 10	1	3	1	3	
10.0 - 20	1	4	6	1	
20.0 - 35	3	6	1	4	2
35.0 +		3	2		2

d. Generate Pareto Data for BPRs across the HLT

Figure 75 illustrates each of the 17 BPRs used in the study and how many times each BPR limited participant Course performance. BPR 17 (Perceptual Integration Capacity-Snowy Pictures) predominated as the limiting BPR for Course performance (N = 11). This subtask required tracking the CDI-H, then make appropriate control inputs to the simulator. BPR 17 demonstrates the capacity to collect pieces of information to make appropriate control inputs. Agreement exists between the reality of the task and these findings.

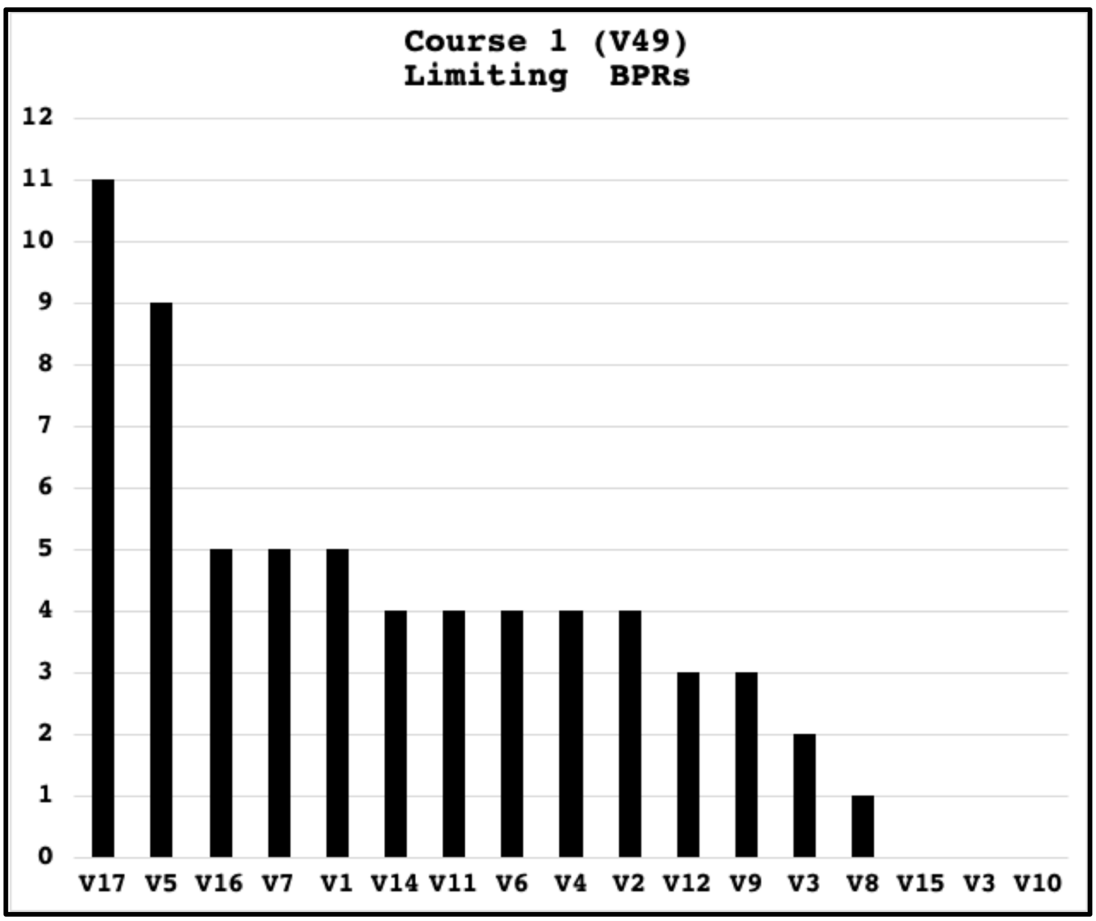


Figure 75. Course 1 Limiting BPRs

e. *Construct Participant R_A Profiles*

Figures 76 and 77 provide a sample of participant R_A profiles including forecasted performance for each BPR. These profiles represent the conditional expectation: E [HLTp | R_A]. Recall, the lowest forecasted performance value determined forecasted performance. For example, participant 1 BPR capacities forecasted a Course performance of 11.65. The highest observed participant performance (86) served as the maximum forecastable value.

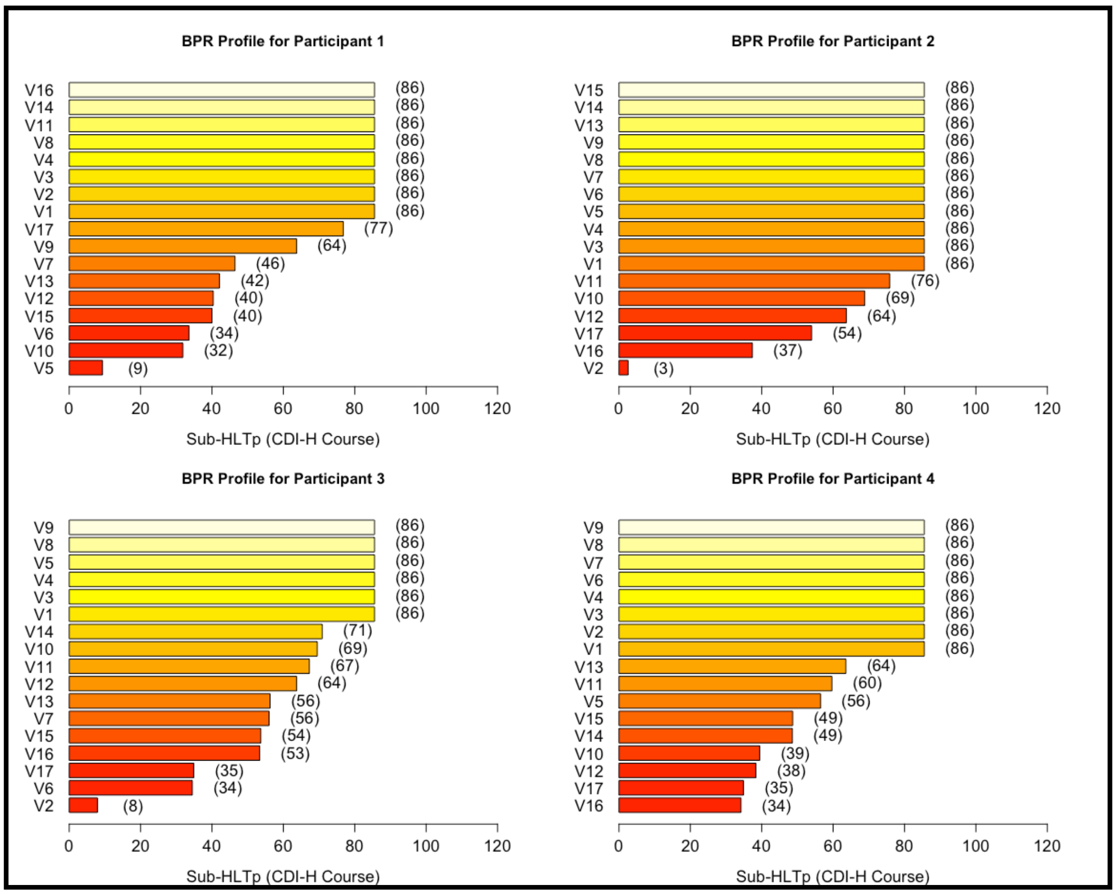


Figure 76. Sample Course R_A Participant Profiles

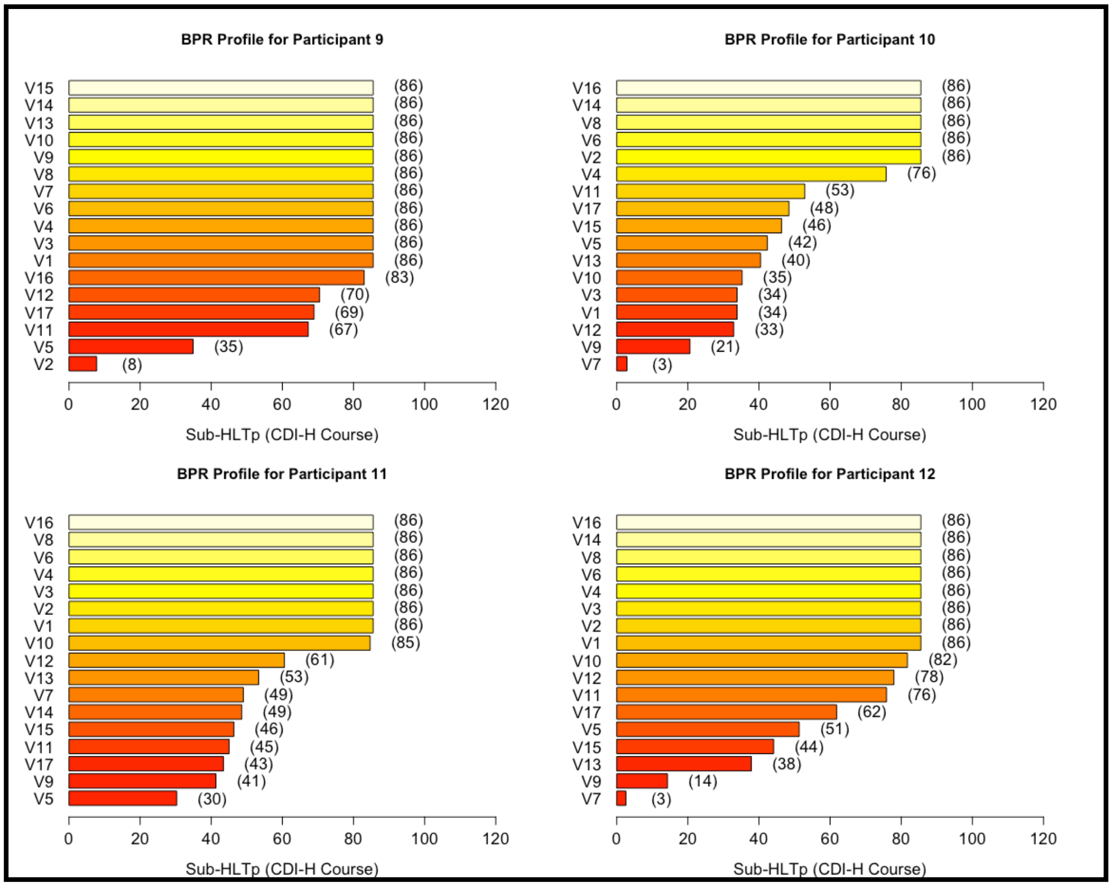


Figure 77. Sample Course R_A Participant Profiles

f. Construct R_D Profile for a Set of HLTp Values

Figure 78 provides a sample of Course R_D profiles. Again, different BPRs predominate at various levels of performance. The significant R_D noted in Figure 78, at a Course performance of 65 demonstrates the delayed resource requirement discussed in Chapter V Section 2. This evidence at different levels of the system hierarchy may prove of high value in supporting SE.

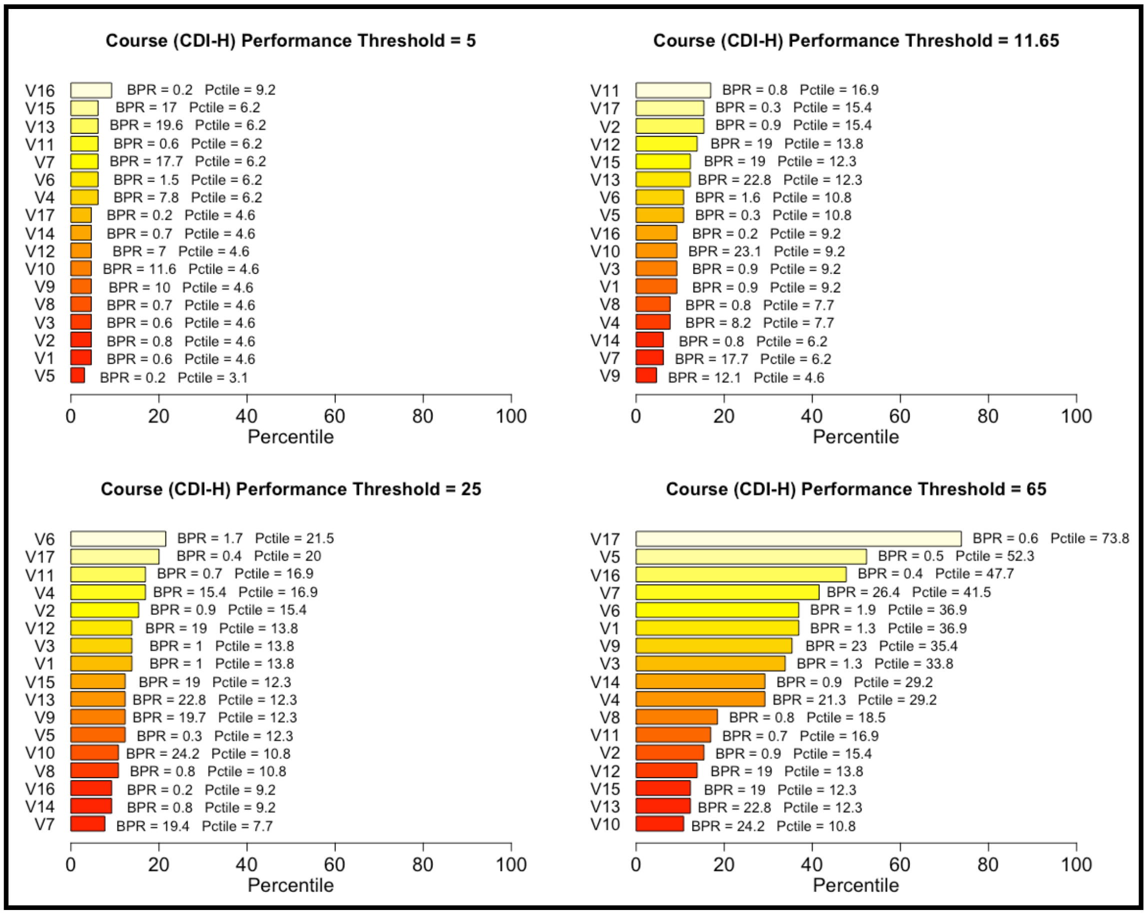


Figure 78. Course 1 without Penalties, Various Performance Levels

In summary, this systemic, iterative, and systematic approach demonstrated again in this section (steps 1–9) is another representative example of MBHSI's value proposition to MBSE. Specifically, this example documents how MBHSI sub-HLT models are built, forecasts are executed, accuracy is measured, and HSI resources (inputs) and performance (outputs) are quantified. The quantification of sub-HLT R_A and R_D, further support establishment of initial constraint values for optimization program formulations discussed in Chapter VII. Additionally, the sub-HLT data outputs from this process produce insights regarding limiting BPRs, predominance of BPR limitation within the sample set, and resource limitations at the individual level. Next, the same evaluation of Glideslope is presented to demonstrate reliability across levels of the HLT hierarchy.

5. Glideslope Sub-HLTp 1 (CDI-V)

Glideslope represents last of the three HLT subtasks. This section also repeats section 5.3.1 procedures using Glideslope. This demonstration further examines the capacity for MBHSI to execute the process at different levels of the task hierarchy. Additionally, this subtask confirms capacity of MBHSI to execute the process across lateral subtasks of an HLT.

a. Plot MBHSI Models and Examine for Errors (from Project I)

Project 1 provides the Glideslope models. The models in step 2 include these models with fitted RDFs.

b. Construct RDFs Along Each BPR in the Model Set

Figures 79 and 80 provide all 17 BPR/Glideslope models with RDFs. The RDFs encapsulate 93% of the data in each model (quantile = .07). These models appeared to demonstrate better agreement when slightly over-fitting the RDFs as discussed in the prior section (Airspeed). However, when overfitting these particular models, Type II errors (under prediction) increase, as expected. Thus, the models demonstrate standard RDF fitting. Again, fitting should be defined by stakeholder priorities. In this case, the RDFs demonstrate a priority of forecast accuracy. The Glideslope models continue to demonstrate agreement with NCRA concepts.

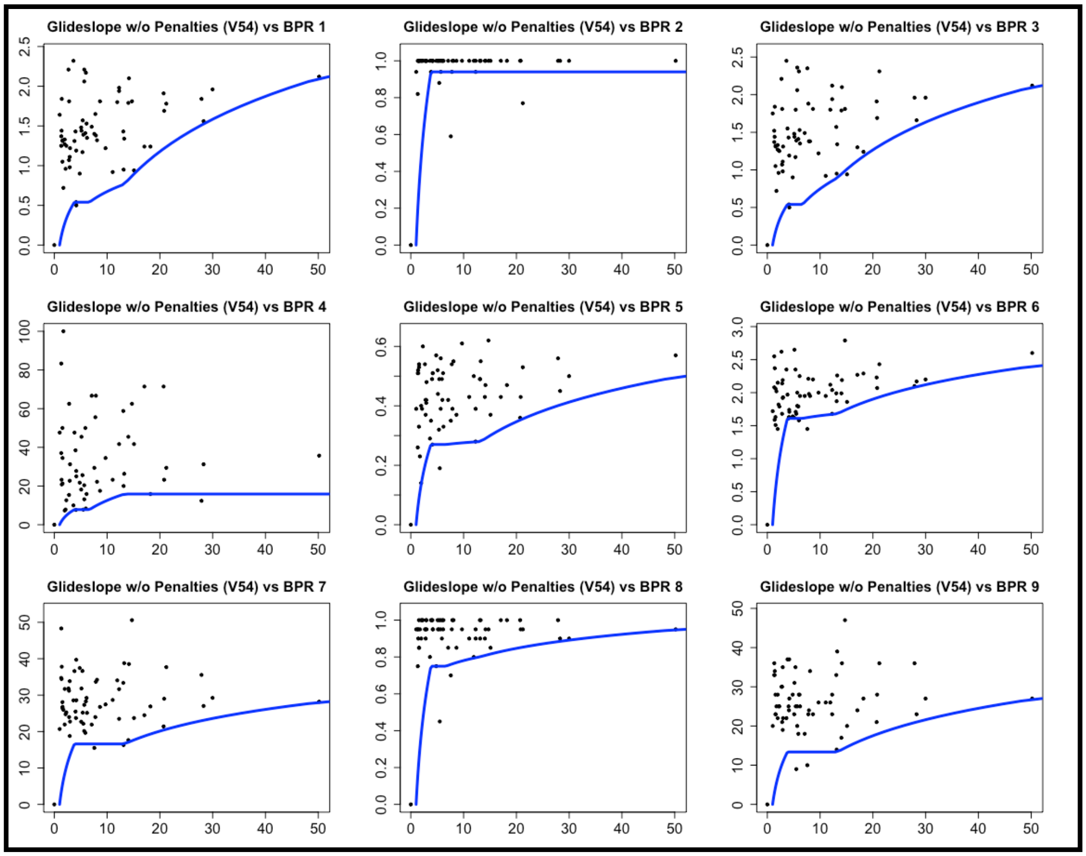


Figure 79. Glideslope without Penalties, RDFs (1–9)

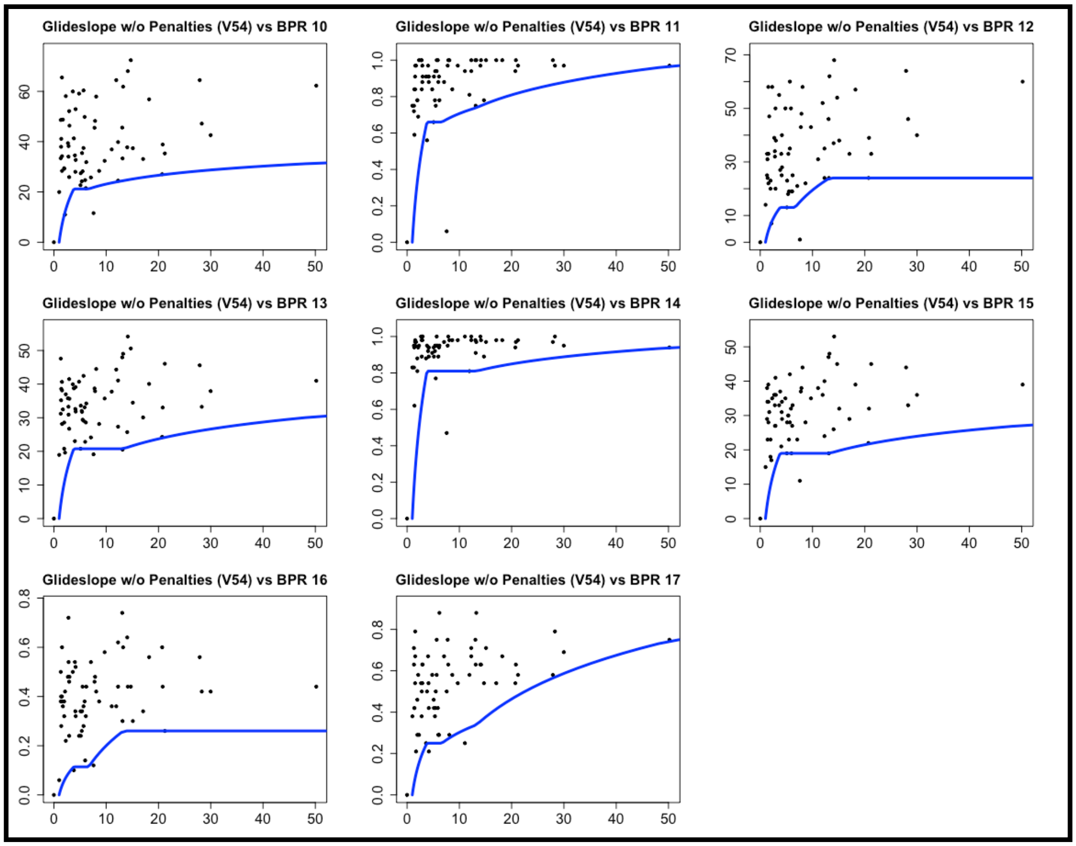


Figure 80. Glideslope without Penalties, RDFs (10–17)

c. *Appraise Difference between Average Predicted HLTp and Actual HLTp Values*

Step 6 documents the outcomes of steps 3–5. Figures 81 and 82 provide a brief review of the actual and predicted data. The boxplot in Figure 81 illustrates actual and predicted distributions. Figure 82 presents the actual vs. predicted data in a scatterplot. Over and under prediction is identified by the red diagonal line. A Spearman's rank-order correlation evaluated the relationship between actual vs. predicted. There was moderate positive (monotonic) correlation between actual and predicted HLTp, which was statistically significant ($r_s = .41$).

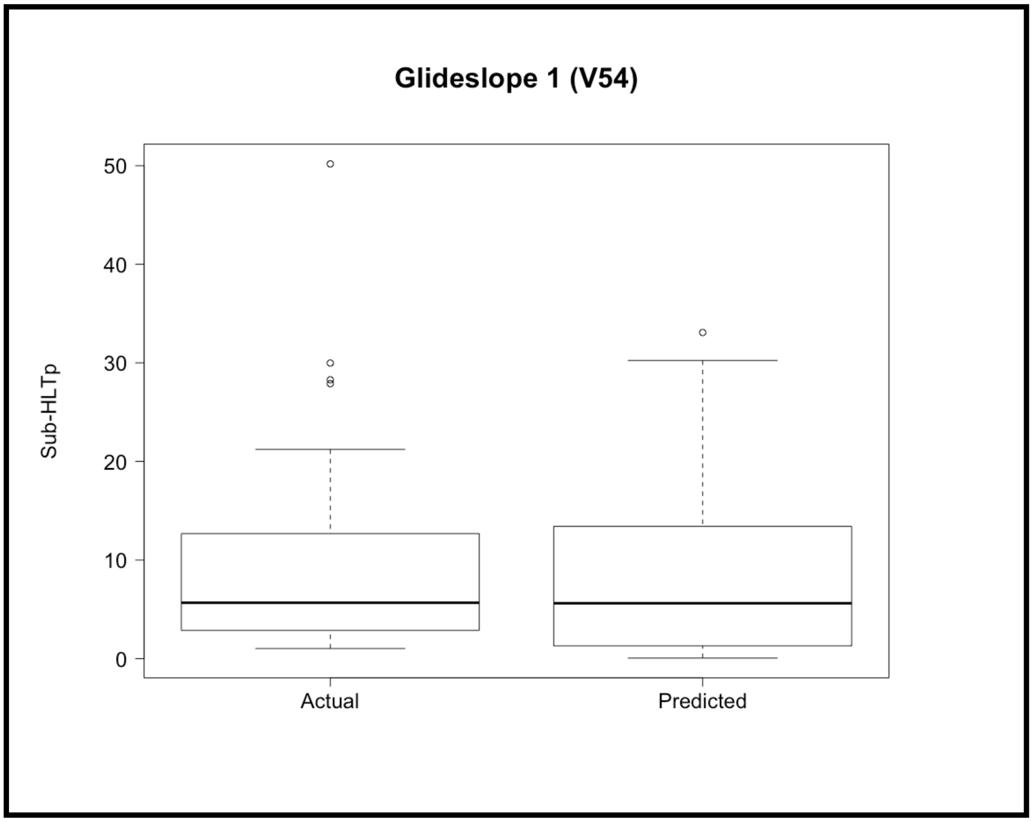


Figure 81. Glideslope 1 Actual vs. Predicted Scores Boxplot

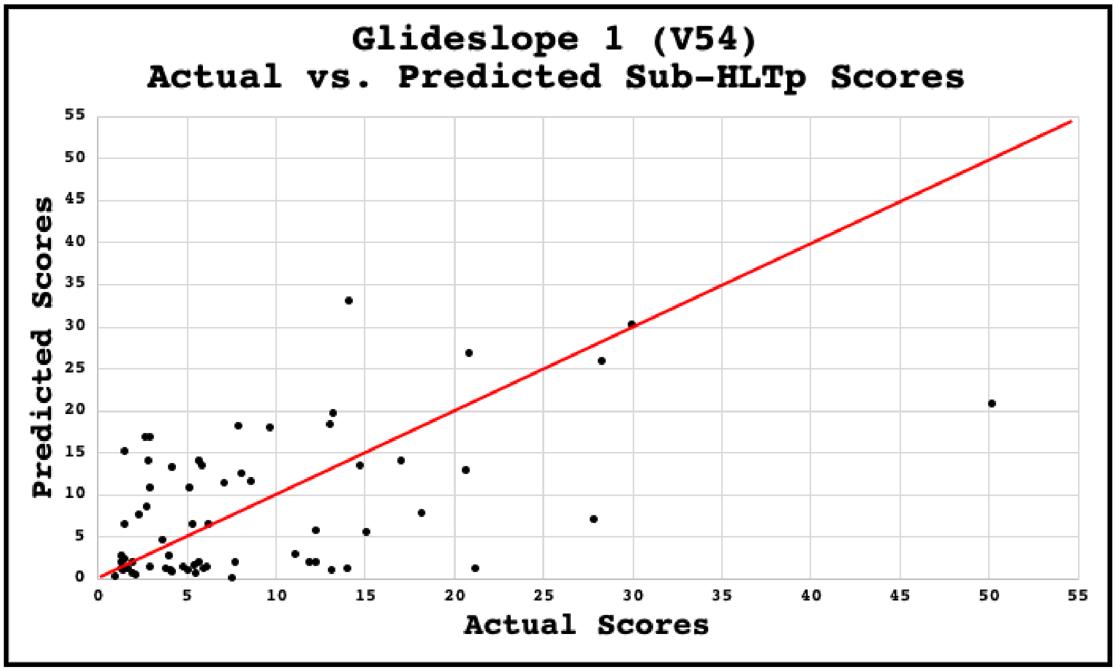


Figure 82. Glideslope 1 Actual vs. Predicted Scores Scatterplot Data (No Scaling)

Table 25 publishes actual median Sub-HLTp, median predicted HLTp, error results, and accuracy using Course observed score (*Mdn* = 5.67). The performance forecast accuracy rate (99%) suggests that the forecast models demonstrate agreement between predicted and actual performance.

Table 25. Glideslope Forecast Accuracy

Glideslope 1 (V54) - Median Scores		
Actual	**Predicted**	**Accuracy**
5.67	5.61	99%

Table 26 provides the model's error rates using the median Glideslope score. Adjustment of the RDF may improve error avoidance but forecast accuracy may degrade. The priorities of the stakeholders determine the proper RDF fit for the system.

Table 26. Glideslope 1 Forecast Results Using the Median

Glideslope 1 (V54) Score = 5.67	Actual Sub-HLTp Score	
Predicted	FALSE	TRUE
FALSE	21	12
TRUE	11	20
True/False Accuacy	**Type I Error Rate**	**Type II Error Rate**
64%	0.34	0.38

Table 27 demonstrates improved Type I (.25) and degraded Type II error forecasts at a score of 10. In total, the model accurately detected 40/64 participant outcomes (accuracy = .63). This evidence further supports the assertion made in the prior section (HLTp) that accuracy regarding errors may be influenced by HLTp level and the RDF formulation. The Type II error results are due to under-prediction of performance at the higher levels of the sub-HLT. The challenging nature of this sub-HLT resulted in fewer high scores thus, the RDF accuracy may be limited with such few data occupying the higher performance areas of the model.

Table 27. Glideslope Type I and II Error Rates with a Score = 10

Glideslope 1 (V54) Score = 10	Actual Sub-HLTp Score	
Predicted	FALSE	TRUE
FALSE	30	14
TRUE	10	10
True/False Accuacy	**Type I Error Rate**	**Type II Error Rate**
63%	0.25	0.58

Table 28 organizes predicted and actual Glideslope 1 scores in a matrix by categories. A Cramer's V was also run to determine strength of the diagonal in a square contingency table. The effect size (ES) was strong (ϕ` = .4) (Cohen, 1988). A Cramer's V was also run using a 2x2 matrix to determine strength of the diagonal using the study median (*Mdn* = 5.67). The ES was moderate ($\phi$` = .25) (Cohen, 1988).

Table 28. Glideslope 1 Actual vs. Predicted Contingency Table

Cramer's V = .4	**Actual Sub-HLTp Score**				
Predicted	< 5	5.0-9.9	10-14.9	15-19.9	20+
< 5	17	8	5	0	1
5.0-9.9	3	2	1	2	1
10-14.9	3	6	1	1	1
15-19.9	3	2	2	0	0
20+	0	0	1	0	4

Table 29 evaluates and documents actual sub-HLT performance vs. absolute forecast error. Actual score categories reflect Table 28 but use smaller categories for absolute error in an effort to gauge how well the model performed across the sub-HLTp continuum. The yellow and green data represent acceptable and desired forecast performance. The majority of participant forecasts fell within the acceptable zones (N = 58/64). Recall the range of actual scores (0-50) observed out of a possible range of 0 -100. Red shaded regions reflect significant overprediction (N = 6/64). These findings suggest limiting BPRs for these particular individuals were not part of the study as demonstrated in the last section.

Table 29. Glideslope Actual vs. Absolute Error

	Actual Sub-HLTp Score				
Absolute Error	< 5	5.0-9.9	10-14.9	15-19.9	20+
< 5	18	10	2	1	2
5.0 - 8.9	5	5	3		2
9.0 - 12.9	2	1	3	2	
13.0 - 16.9	3		1		
17.0 +			1		3

d. Generate Pareto Data for BPRs across the HLT

Figure 83 illustrates each of the seventeen BPRs used in the study and how many times each BPR limited participant Glideslope performance. BPR 6 (Multi-Choice Reaction Speed) predominated as the limiting BPR for Glideslope performance (N = 9). BPR 17 (Perceptual Integration Capacity-Snowy Pictures) and BPR 12 (ANAM-Spatial Orientation Throughput) accounted for 14 participant-limiting BPRs. This subtask required tracking the CDI, then making appropriate control inputs to the simulator. Agreement exists between the reality of the task and these findings. BPRs V15, V13, V11 did not predominate as a limiting BPR for Glideslope.

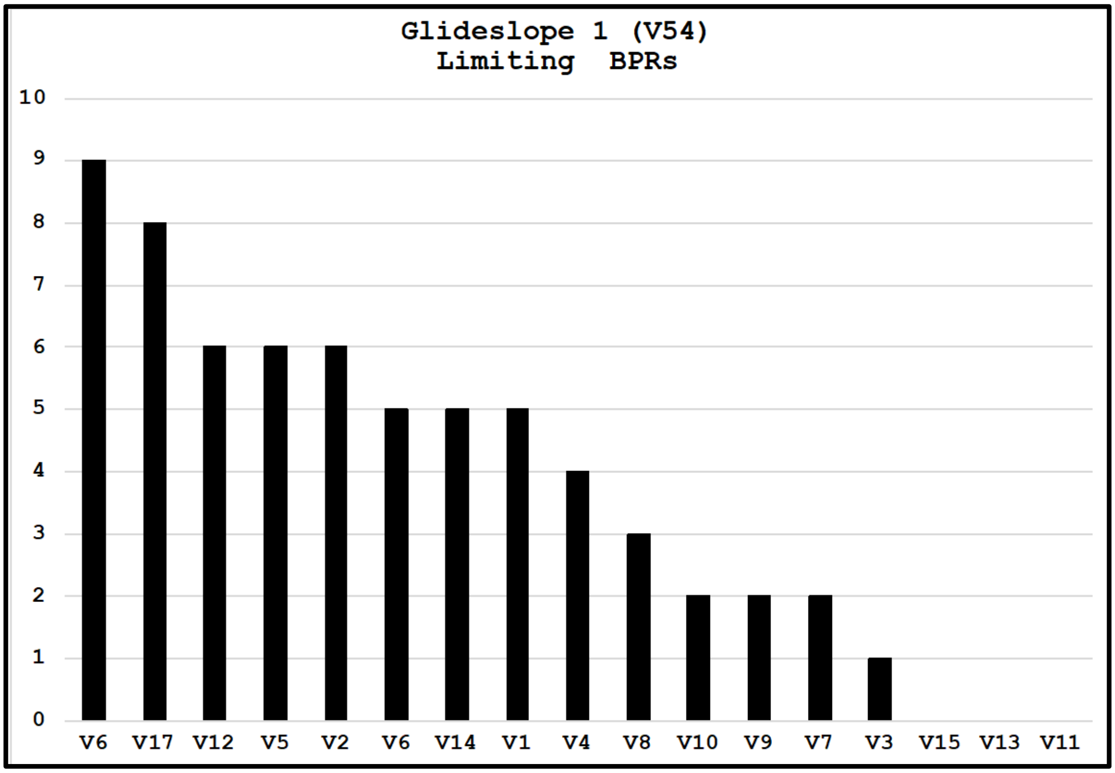


Figure 83. Glideslope Limiting BPRs

e. Construct Participant R_A Profiles

Figures 84 and 85 provide a sample of participant R_A profiles including forecasted performance for each BPR. These profiles represent the conditional expectation: E [HLTp | R_A]. The lowest forecasted performance value determines forecasted performance. For example, participant 1 BPR capacities forecasted Glideslope performance of 5.67. The highest observed participant performance (50) served as the maximum forecastable value. Continued variation persists along not only capacities, but also the order of BPRs on the y axis by participant.

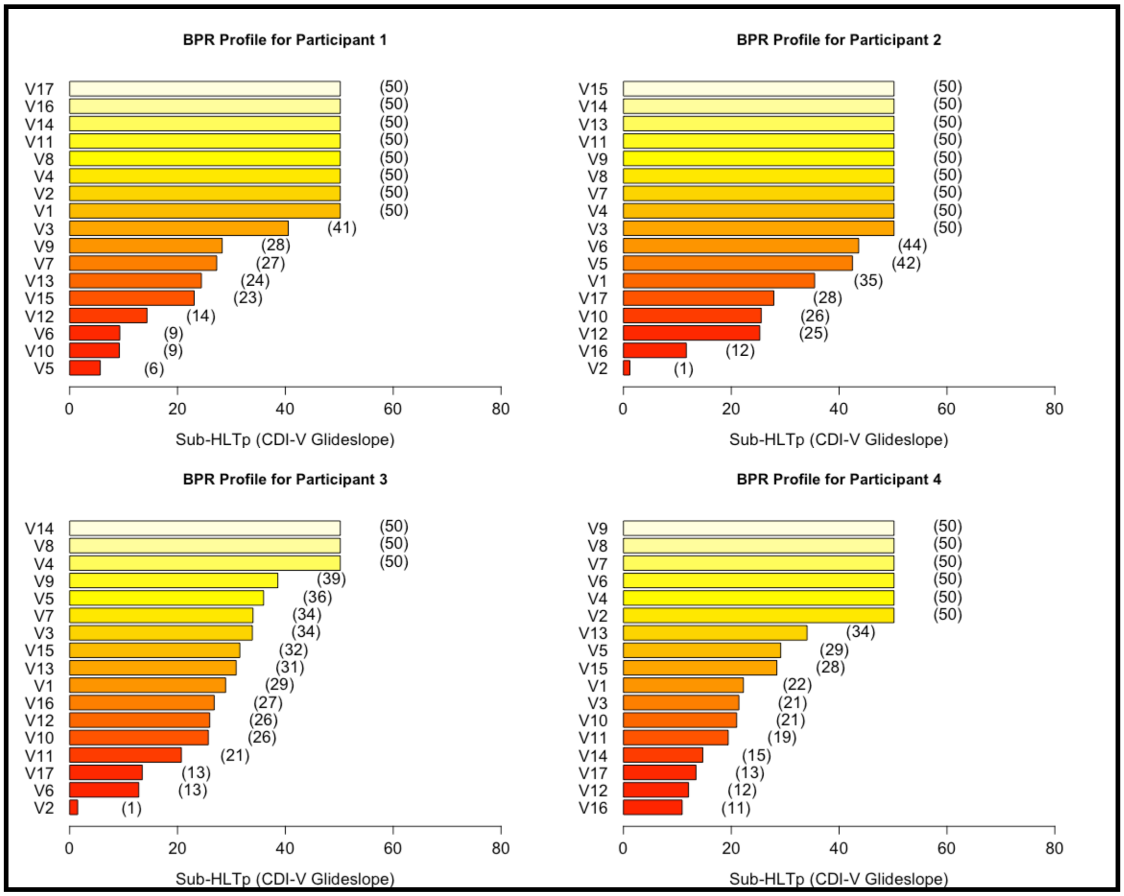


Figure 84. Sample Glideslope R_A Participant Profiles (1–4)

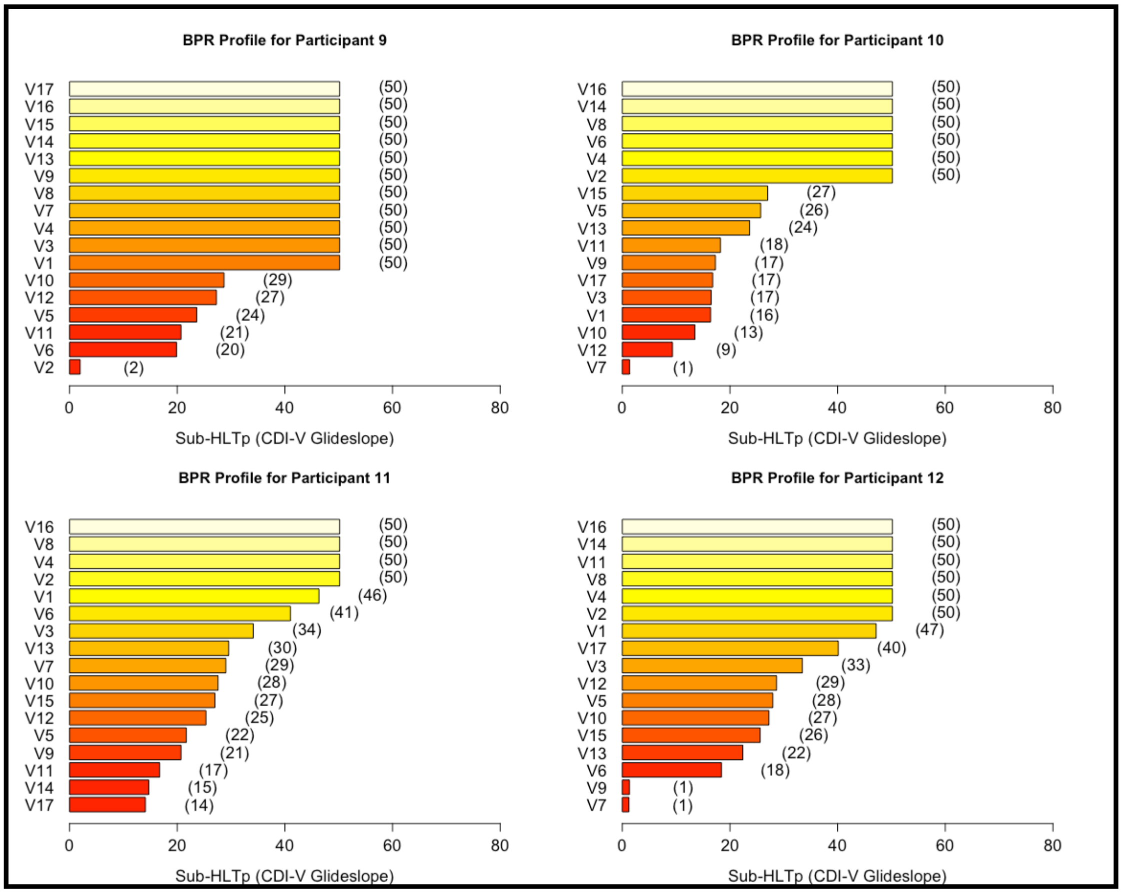


Figure 85. Sample Glideslope R_A Participant Profiles (9–12)

f. Construct R_D Profile for a Set of HLTp Values

Figure 86 provides a sample of Glideslope R_D profiles. These profiles represent the conditional expectation: E [R_A | HLTp]. Each profile lists forecasted threshold-levels of BPRs by HLTp. Each figure documents the minimum capacity expected across each of the seventeen BPRs for a given level of Glideslope performance. Again, different BPRs predominate at various levels of performance. This additional evidence at lateral levels of the system hierarchy may prove of high value in supporting SE.

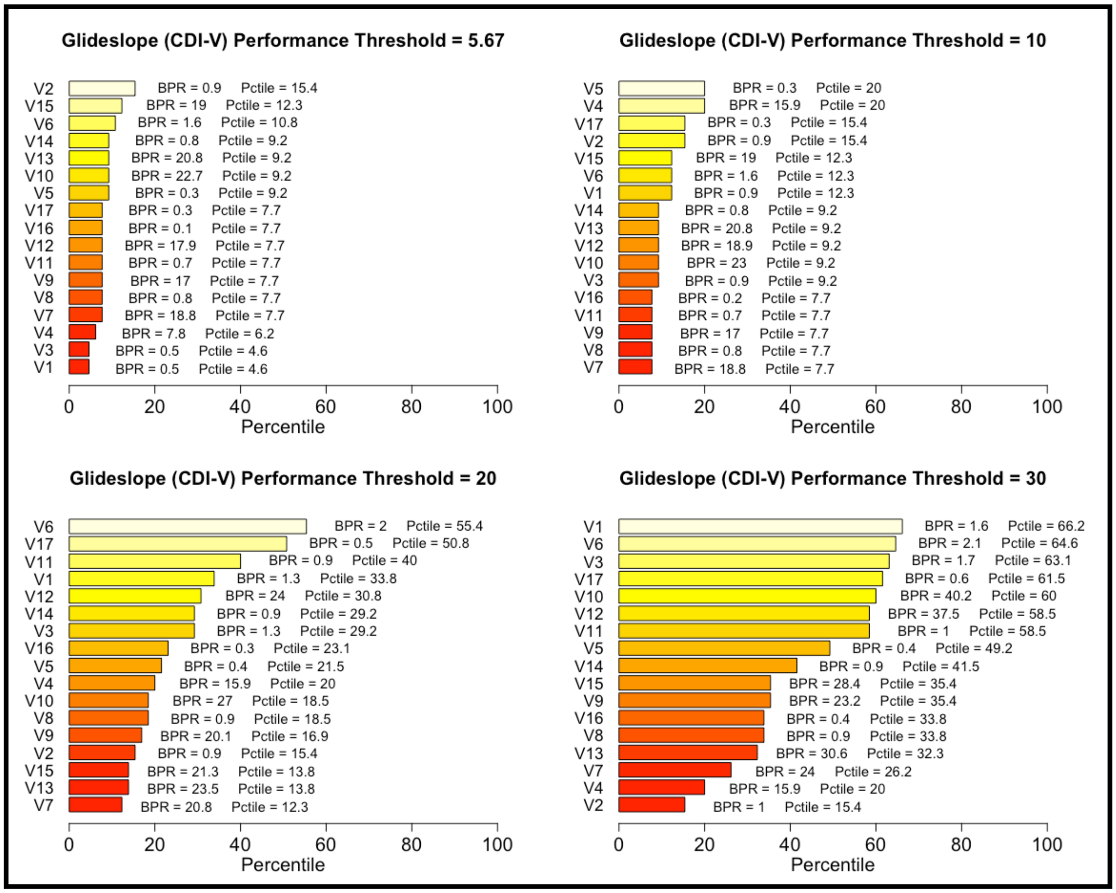


Figure 86. Glideslope R_D Profiles

In summary, this systemic, iterative and systematic approach demonstrated again in this section (steps 1–9) a third sub-HLT level as another representative example of MBHSI's value proposition to MBSE. Specifically, this example documents how MBHSI sub-HLT models are built, forecasts are executed, accuracy is measured, and HSI resources (inputs) and performance (outputs) are quantified. The quantification of sub-HLT R_A and R_D, further support establishment of initial constraint values for optimization program formulations discussed in Chapter VII. Additionally, the sub-HLT data outputs from this process produce insights regarding limiting BPRs, predominance of BPR limitation within the sample set, and resource limitations at the individual level.

C. DISCUSSION

The MBHSI RDFs developed in this study demonstrate quantitative relationships between HSI resources (BPRs) and system-level performance. The RDFs suggest agreement between forecasted performance at two different levels of the task hierarchy. They also suggest agreement laterally within the sub-HLTs of the study. Various fitting techniques demonstrate the capacity of the RDF to forecast performance that might accommodate stakeholder priorities. For example, a priority of accuracy demonstrated slightly higher error rates. While over-fitting is observed on some RDFs, improved accuracy is also achieved. A limitation of the RDF is noted when few data points exist at particular levels of performance. The leverage of the less dense data in the models might inappropriately influence the RDF. For example, a single data point might demonstrate inappropriate leverage on the RDF resulting in inaccurate performance forecasts. As stated, stakeholder priorities and defined performance thresholds provide excellent boundary conditions for MBHSI. Additionally, because GSPT is not correlation based, data resolution in the models should be the focus when determining data quantity. If insufficient resolution at or near a desired HLTp is noted, then additional data should be obtained around that value to increase confidence in the RDF placement.

Percent differences suggest agreement with forecasted and observed performance across all four model sets. In many cases, where relatively large disagreements are noted, a prevalence of over prediction is observed. Over predictions noted using NCRA suggest limiting BPRs exist for participants which were not included in the study.

In general, limiting BPRs demonstrate agreement with the HLT or sub-task. For example, multi-limb coordination speed predominates as the limiting BPR for Airspeed. However, the data suggest other BPRs limited performance for many participants in the study. These findings support the notion that individual systems (humans) might be limited by different resources when accomplishing the same task. The capacity to identify which resources limit each participant appears to be a valuable attribute of MBHSI.

The resource profiles documented in this study provide significant insight into NCRA. These profiles quantitatively capture individual capacities across the measured

BPRs and their respective forecasts for each evaluated model set. Additionally, the resource demand functions capture powerful data quantifying what resources predominate at various levels of performance for the same task. Of particular interest is the variance in the order of necessary resources at each level of performance. Without these insights, forecasting the trade space appears impossible. These data demonstrate that performance levels matter. Performance expectations will drive a specific requirement for certain resources and their values.

In closing, this study realized the conceptual outputs of NCRA to enable MBHSI. The establishment of the relationships between HSI resources (BPRs) and system-level performance (HLTp) demonstrate a human system model. This requirement appears to be satisfied based on the documented evidence. This improved capacity of HSI suggests a complementary model-based approach to MBSE appears promising. Therefore, both stated purposes for this study—the realization of NCRA and the achievement of mapped HFRs—were achieved. The outputs of this study provide quantitative data in support of Project III.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. RESEARCH PROJECT III: MBHSI TRADE SPACE

We can't solve problems by using the same kind of thinking
we used when we created them.

—Albert Einstein

A. PROBLEM STATEMENT

DoDI 5000.02PR (draft) requires HSI to "integrate human considerations into system design to optimize total system performance and minimize life-cycle costs" (Department of Defense, 2020, p. 1). If TSP reflects how well a system has integrated the human, then HSI's capacity to integrate human considerations directly impacts DOD's TSP. According to the USAFSAB, "the critical function of HSI [has] become insufficient" (U.S. Air Force, 2012). HSI lacks a complementary model-based approach to integrate the human system into MBSE system architectures effectively. In order to support MBSE, HSI must develop complementary methods for modeling, forecasting, quantifying HSI trade space, and optimization. This research project suggests a theoretically-based process to articulate a small part of the HSI trade space. This process of resource manipulation and quantification, including posttest performance outcomes, supports value-based design insights. These insights may increase the capacity for HSI to address the trade space and optimization challenges that continue to plague the DOD, in terms of HSI.

B. OBJECTIVE

Project III serves two primary purposes. The first purpose is to quantify the effects of PCE (R_A) expansion (i.e., training) on HLTp. This increase in R_A changes the shape and/ or volume of the N-dimensional PCE while maintaining the HLT. The second purpose is to quantify the effects reducing HLT R_D on HLTp (i.e., positive HFE). This reduction in R_D targeted the HLT while maintaining the PCE. These methods support evidence-based insights to MBSE by bringing this quantification method to the complex HSI trade space.

The secondary purpose of Project III is to satisfy the fifth MBHSI-FR, *quantitatively articulate the HSI trade space*. The criteria for success is the quantification of study constraints following relaxation and restriction. Figure 87 illustrates the MBHSI-FR mapped to Project III.

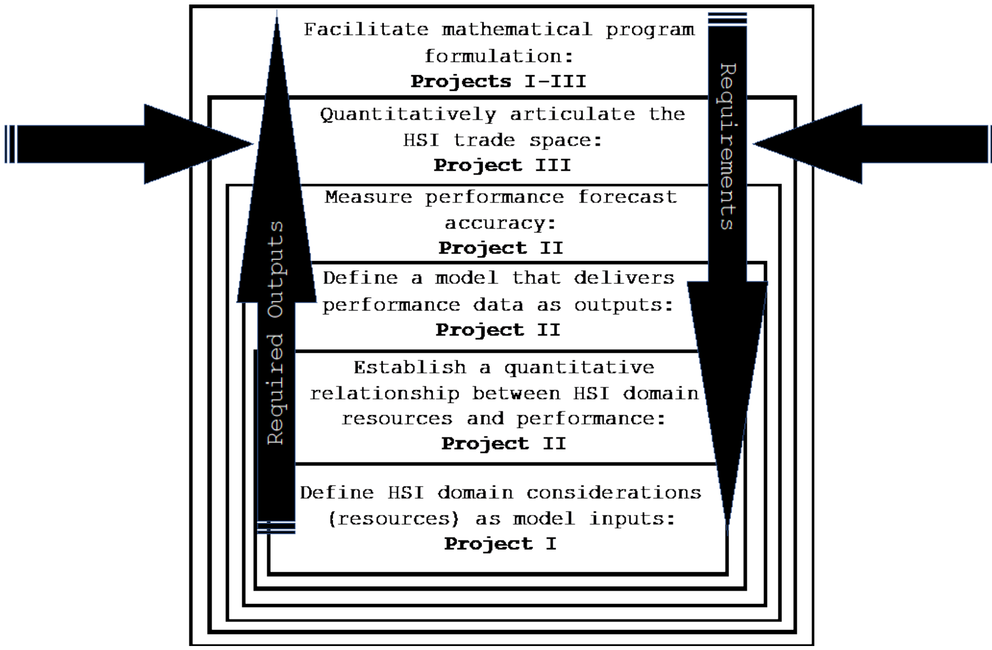


Figure 87. Project III as Mapped to MBHSI-FRs

C. PROJECT III OVERVIEW

Project III independently manipulates R_A and R_D to generate system performance data. Quantitative outputs from training (R_A), HFE (R_D), and posttest HLTp quantify a select portion of the HSI trade space for demonstration. The outcomes appear to complement MBHSI's diagnostic capacity, informing an individual or systemic prescriptive response. Specifically, if MBHSI can diagnose limiting BPRs, then evidence-based changes can target specific R_A and/or R_D to resolve individual or systemic performance issues. A roadmap for Project III illustrates this process in Figure 88.

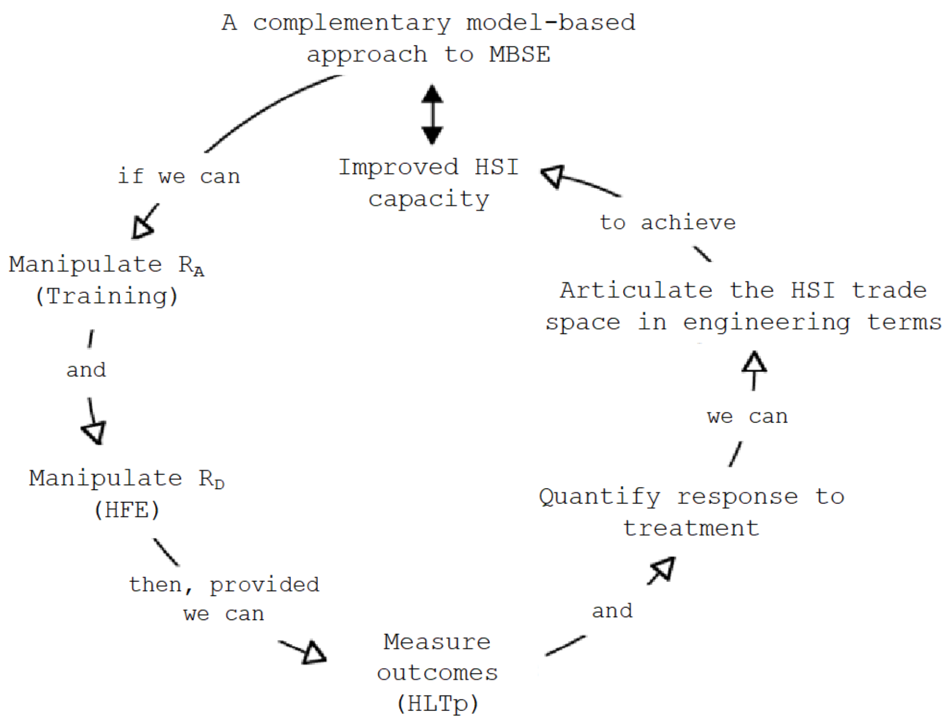


Figure 88. Project III Roadmap. Adapted from Hitchens (1992).

D. HYPOTHESIS

An HSI trade space can be assessed quantitatively by manipulating R_A or R_D and measuring the impact on HLTp.

E. METHODOLOGY

The project moves on from model development and forecasting to demonstrate hypothesis testing that is possible with GSPT and MBHSI. Specifically, this project manipulates the PCE by providing training and leaving the HLT fixed. Then, the HLT is manipulated by introducing the autopilot feature in the simulation and leaving the PCE stable. The goal is to quantify the effects on HLTp. This section describes Project III's sample, research design, variables, instruments, procedures, analytics, and results. Recall Track B participants included those that received training. Track C.1 participants had the sub-task Course automated. Track C.2 participants had the sub-task Glideslope automated.

F. SAMPLE

Project III used data from all sixty-four participants. Track A was the control group (n = 40). This project focused on the ten participants who received professional video-based training (Track B) and the fourteen participants who received a modified HLT (Tracks C.1 and C.2). Track C.1 (n = 7). Track C.2 (n = 7).

G. RESEARCH DESIGN AND VARIABLES

Project III was a mixed-design. Project III variables:

- Track B: The IV was training level (provided or not provided). The DVs are knowledge tests and HLTp scores.

- Tracks C.1 and C.2: The IV was automation (provided or not). The DV was the HLTp scores.

H. INSTRUMENTS

Project III Track B (Training) used professionally developed flight training videos produced by Jeppesen Sanderson Inc. for each participant (n= 10).

- Jeppesen Sanderson Inc. (2006a). Guided Flight Discovery: Private Pilot. Englewood, CO: Jeppesen Sanderson, Inc. (Jeppesen Private Pilot DVD1, 0:00-8:00, n.d.).

- Jeppesen Sanderson Inc. (2006b). Guided flight discovery: Instrument/ Commercial video series. Englewood, CO: Jeppesen Sanderson, Inc. (Jeppesen Instrument Commercial DVD2, 1:57:17-2:14:00, n.d.).

Project III Track C (HFE) used X-Plane 11 autopilot functions to automate either Course (n = 7) or Glideslope (n = 7).

I. PROCEDURES

Track B participants watched both Jeppesen Sanderson Inc. training videos back-to-back following completion of HLT 1. The training lasted approximately twenty-five minutes. Upon completion of the video training, each of the Track B participants completed the same knowledge tests a second time. The participants had unlimited time to complete the tests. Upon completion, the participants (n = 10) completed HLT 2. Following completion of HLT 2, each participant received a certificate of appreciation concluding the study.

Tack C participants completed a modified HLT 2. Track C.1 participants (n = 7) completed HLT 2 with the Course automated during the entire ILS approach. Track C.2 participants (n= 7) completed HLT 2 with the Glideslope (CDI-V) automated. Upon completion of HLT 2, each participant received a certificate of appreciation concluding the study.

J. ANALYSIS

The Track B IV was training using a within- subjects pre-posttest design. Track B DVs included post-training knowledge test and HLTp 2 scores. Track A participant BPR and HLTp data served as the control group. Track B IV/DV data were compared to the control group data using a between- subjects pre-posttest design.

The Track C IVs included modified HLTs (automation) using a within -subjects pre-posttest design. Track C DVs included HLTp 2 scores. Track A participant HLTp 1 and HLTp 2 scores served as the control group. Track C DV data were compared to the control group data using a between- subjects pre-posttest design.

Output data (HLTp scores) failed to meet assumptions and conditions for parametric evaluation. Therefore, Project III study data used nonparametric statistical evaluation techniques. Wilcoxon Rank Sum Tests (WRST) and median tests compare variables in the study. All analyses were conducted using R statistical programming language using a 0.05 level of significance. Finally, GSPT/NCRA methods used in Projects I and II data establish the trade space between Project III resource variables and performance.

1. Track B - R_A Expansion: Training

a. Results

The reader will recall the chapter objective stated the first purpose was to quantify the effects of PCE (R_A) expansion (i.e., training) on HLTp. This increase in R_A changes the shape and/or volume of the N-dimensional PCE while maintaining the HLT. Table 30 documents all of Track B comparative analysis results. Overall, R_A expansion (training) significantly improved participant knowledge BPRs (WRST #6 and #7). Flight and aircraft instrument knowledge median scores improved forty-two and 26 percent respectively. These results suggest that the training was sixteen percent more effective at increasing flight knowledge vs. aircraft instrument knowledge. While a lower *p*-value was noted regarding Track B HLTp 2 scores as compared to Track A (control), they were not statistically significant (WRST #5 and #2). The Track B sample size (n = 10) may be a threat to validity; however, the results do suggest that it is possible to manipulate R_A and measure the impact on HLTp. Appendix M provides a comprehensive set of boxplots of this particular data set.

Table 30. Tack B Comparative Analysis Results

WRST #	Track	Variable Score	Median	Wilcoxson Rank Sums Test
1	A	Flight Knowledge Pre-test	0.29	n/a
	A	Instrument Pre-test	0.33	
2	A	HLTp 1	271.68	$p = 0.231$
	A	HLTp 2	705.84	
3	A	HLTp 1	271.68	$p = 0.070$
	B	HLTp 1	91.04	
4	A	HLTp 2	705.84	$p = 0.952$
	B	HLTp 2	465.3	
5	B	HLTp 1	91.04	$p = 0.089$
	B	HLTp 2	465.3	
6	B	Flight Knowledge Pre-test	0.29	$p = 0.008$
	B	Flight Knowledge Post-test	0.72	
7	B	Instrument Pre-test	0.33	$p = 0.010$
	B	Instrument Post-test	0.59	

b. Discussion

During model building in Project I, the knowledge BPRs did not reflect desired characteristics of GSPT. This was expected because of the novice nature of the participants and the format of the test being primarily multiple-choice. This determination suggested the two knowledge BPRs were not limiting, consistent with the WRST findings. Two critical insights emerge:

1. Knowledge BPRs were not part of univariate HLTp scoring rubric. Unless knowledge capacity is part of the HLTp score, it may be difficult to detect knowledge as a limiting BPR or its effect on HLTp. Knowledge did not appear to influence HLTp scores in this study. This outcome could also be due to poor training. Future MBHSI work should consider these important aspects.
2. If training seeks to improve HLTp, then training should be directed toward limiting BPRs. Because knowledge BPRs did not appear to be limiting in this study, training may be more effective if directed towards actual limiting BPRs. If improved performance is desired, then the R_D profiles generated in Project II should inform the analysis portion of the instructional system design (ISD) process.

These insights suggest that integrating knowledge capacity into HLTp scoring rubrics and/or targeting training towards individual or population limiting BPRs might result in improved training designs and performance outcomes. Figure 89 illustrates actual limiting BPR R_D profiles spanning four different levels of HLTp. These results also demonstrate the importance of defining performance level prior to engaging in ISD. For example, if an HLTp score of 282 is determined to be satisfactory, then ISD efforts might focus on the predominating BPRs (V11, V3, V2 etc.). Alternatively, if an HLTp score of 5,000 is desired, then targeting BPRs V17, V10, V3, etc., might result in more effective and efficient training. The central idea is to compare R_D to individual R_A for insight into an individualized training plan. Additionally, these data may positively support personnel alignment with the system and could inform system designers where excessive training burden exists to be designed out of the system.

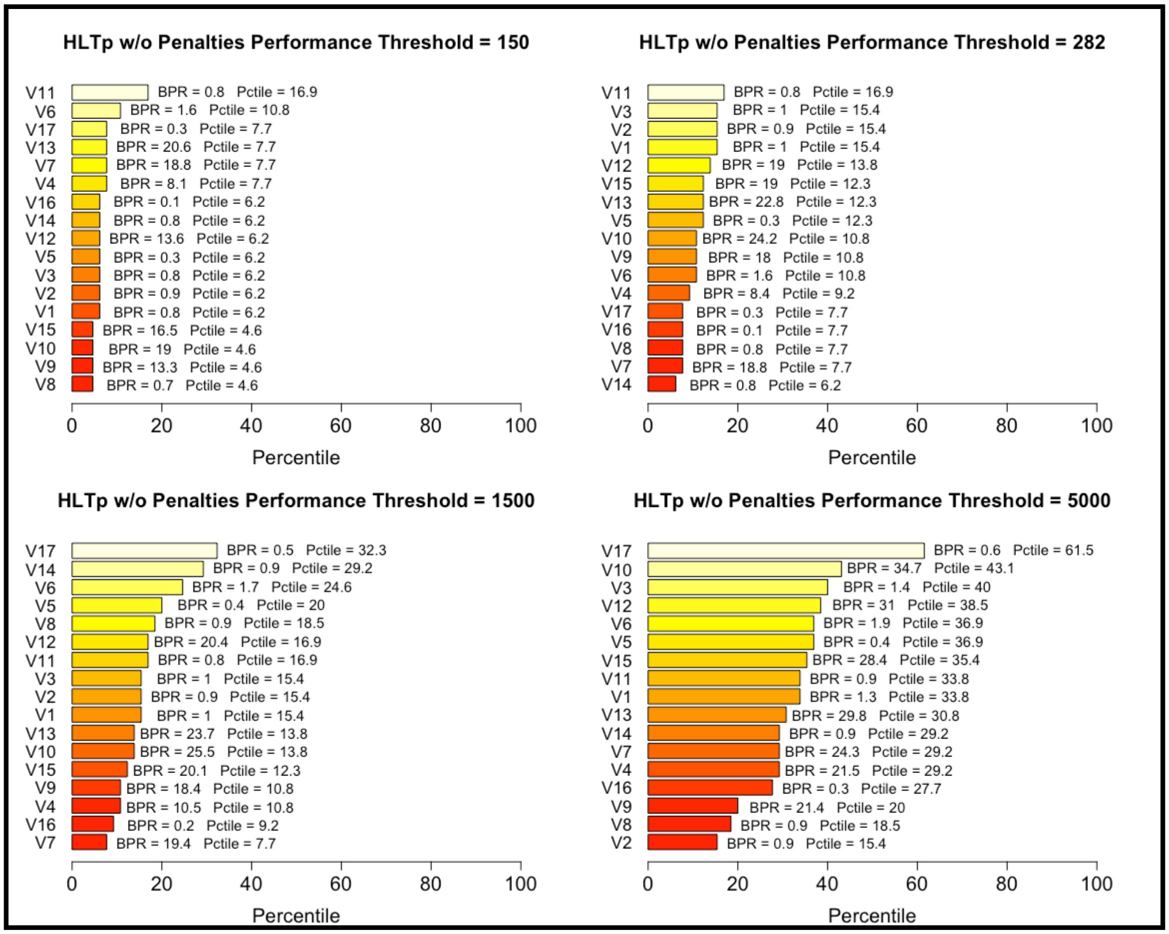


Figure 89. Sample R_D Profiles at Various HLTp Scores

A training requirement noted in DoDI 5000.02PR *draft* specifically instructs "individual" training for operators (p. 7). Designing individual training plans requires relevant baseline resource capacities to be compared to relevant R_D at specified performance levels. The gaps between limiting R_A and R_D articulate potential training targets of opportunity. MBHSI individual BPR profiles shown in Figure 90 articulate R_A. These resources may provide insight regarding individual training baselines. These data suggest that individualized training plans would vary significantly even while targeting the same HLTp. These data also suggest that one-size-fits-all training plans may not be effective nor efficient. Additionally, the amount of R_A needed before another BPR will dominate as limiting demonstrates another key feature of the R_A profiles. This methodology may prove valuable because evidence-based training decisions should guide individual training plans. For example, participant 1 might benefit more from V5 training vs. participant 2, assuming V5 is a malleable resource. If V5 is non-malleable, then these data inform personnel selection criterion and/or system design characteristics.

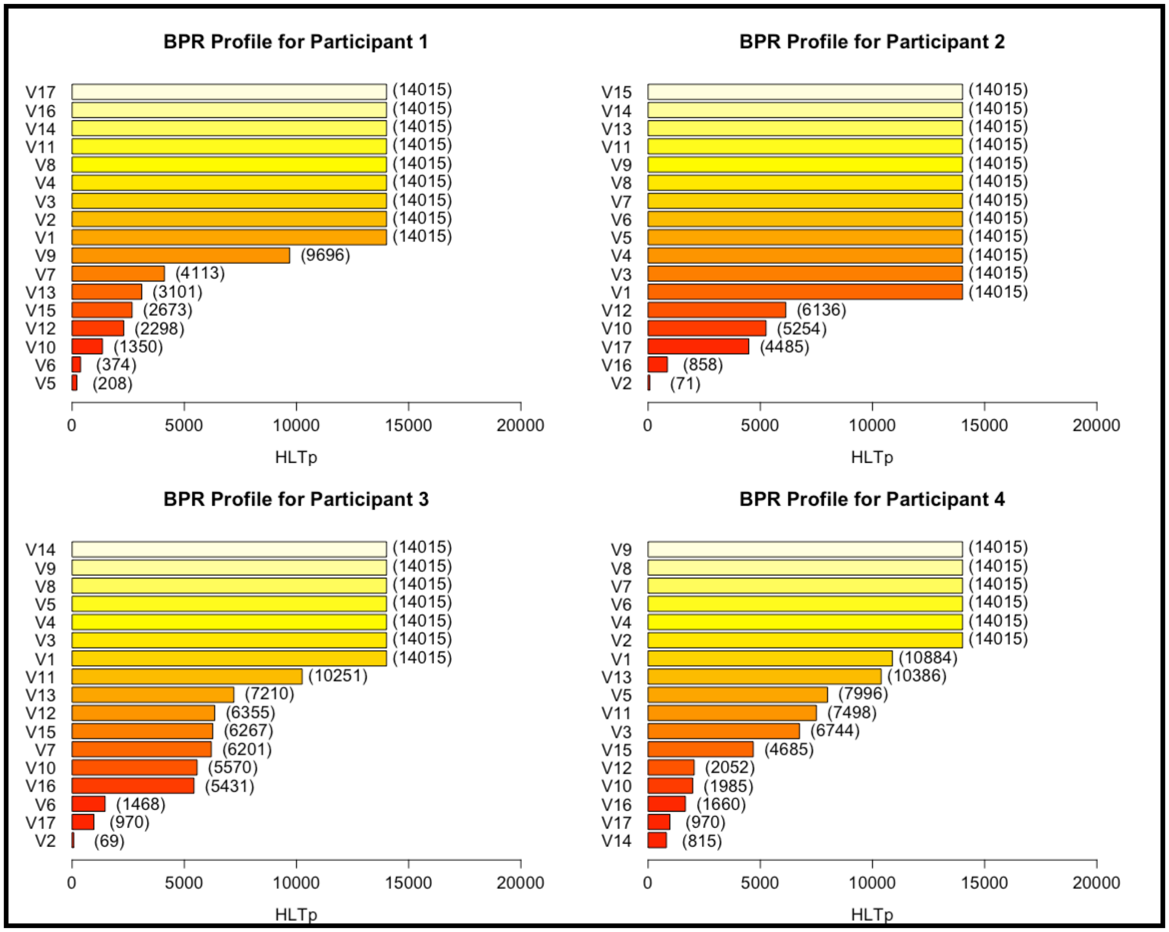


Figure 90. Sample R_A Profiles

2. Tracks C.1 and C.2 - R_D Reduction: HFE

a. Results

The reader will recall the chapter objective stated the second purpose was to quantify the effects of reducing HLT R_D on HLTp (i.e., positive HFE). This reduction in R_D changed the HLT while maintaining the PCE. Overall, automating the Course (Track C.1) and Glideslope (Track C.2) sub-HLTs demonstrated significant differences according to the WRST in HLTp 2 scores (WRST #9 and #12). The data suggest that automating the flight Course (CDI-H) in this HLT resulted in a net *median* HLTp increase of 21,112.38 when compared to Track A (WRST #9). The data also suggest the automation of Glideslope (CDI-V) in this HLT resulted in a net *median* HLTp increase of 41,105.73 as compared to Track A (WRST #12). Therefore, automating Glideslope appears to result in greater HLTp improvement as shown in Table 31.

Table 31. Tack C.1 and C.2 Comparative Analysis Results

WRST #	Track	Variable Score	Median	Wilcoxson Rank Sums Test
8	A	HLTp 1	271.68	p = >.999
	C.1	HLTp 1	267.22	
9	A	HLTp 2	705.84	p = 0.001
	C.1	HLTp 2	21,813.76	
10	C.1	HLTp 1	267.22	p = 0.001
	C.1	HLTp 2	21,813.76	
11	A	HLTp 1	271.68	p = 0.026
	C.2	HLTp 1	2,642.71	
12	A	HLTp 2	705.84	p = < .001
	C.2	HLTp 2	44,182.60	
13	C.2	HLTp 1	2,642.71	p = 0.002
	C.2	HLTp 2	44,182.60	
14	A	HLTp 1	271.68	p = 0.231
	A	HLTp 2	705.84	

Performance forecasts were also computed up one level in the HLT hierarchy, at sub-HLT (Airspeed, Course, and Glideslope). Figure 91 provides the same participant sample of individual performance forecasts at the sub-HLT level. In this example, Course was forecasted to limit HLTp for participants 2, 3, and 4. Airspeed was forecasted to limit HLTp for participant 1. The reader will note that the limiting BPR over predicts HLTp when compared to using the 17 study BPRs. In this case, participants 1, 2, and 3 sub-HLTp forecasts are higher than when forecasted using the BPRs. This is expected due to the limited number of BPRs at play (3) as discussed in Project II. Actual vs. predicted scores at this level were not as accurate (88%) across the study set (N = 64) at the sub-HLT level (281.97 vs. 314.66) comparted to the 97% accuracy noted at the BPR level in the study

(281.97 vs. 289.15). This new perspective of the HLT provides quantitative insights regarding sub-tasks within an HLT to support R_A and R_D decisions within a system. For example, these data might inform individualized training plans. In this case, an effective and efficient training plan for participant 1 might focus heavily on maintaining Airspeed (V44) in the aircraft whereas participant 2 might benefit more from a focus on maintaining Course (V49). Evaluation of these four participants might also suggest that a design focus on Course (V49) might improve HLTp more efficiently than the other two sub-HLTs. However, Glideslope (V54) dominated across the study sample as the limiting sub-HLT compared to Course (V49) as noted by the lower median score (5.76 vs. 11.65).

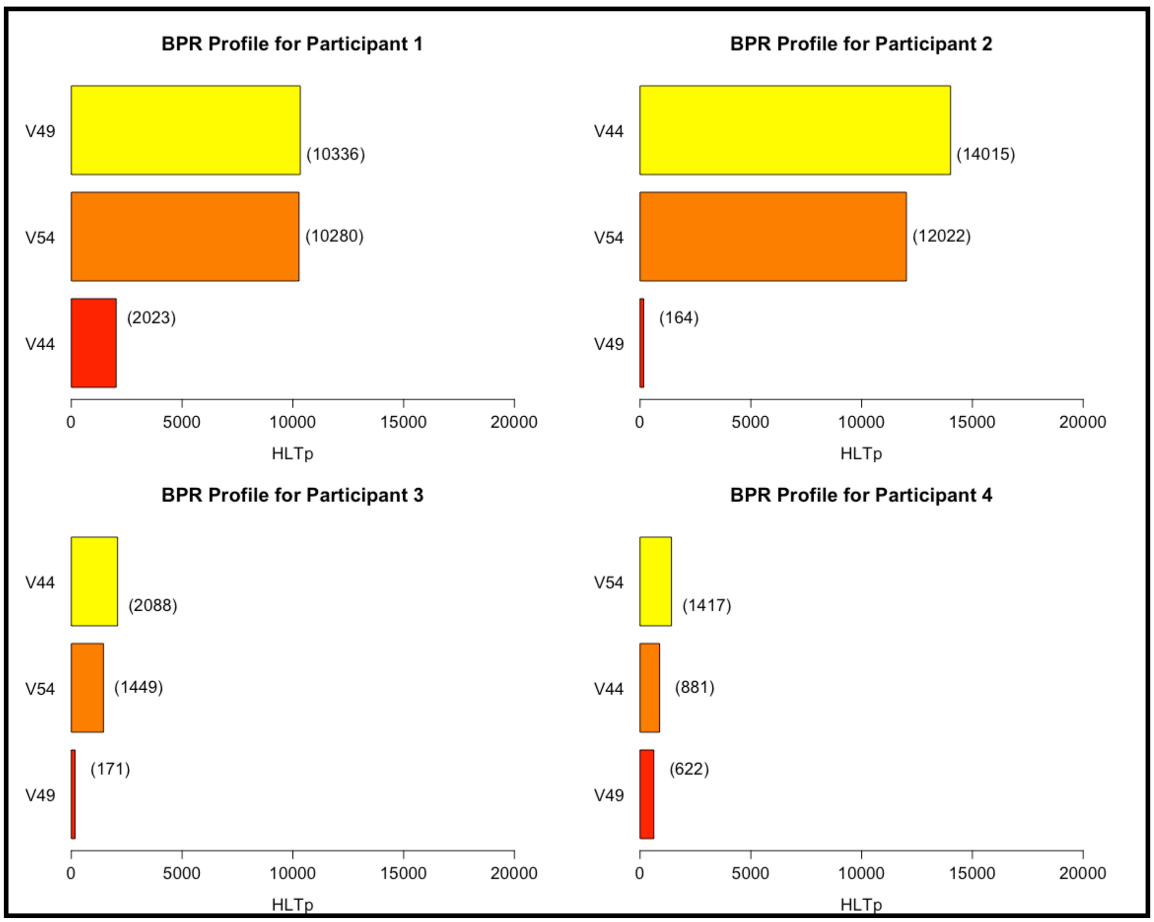


Figure 91. Sub-HLTp Forecasts

Finally, individual BPRs vs. Course and Glideslope in Figures 92 and 93 also articulate MBHSI trade space at the BPR-level by individual. These data suggest quantification of R_A might support personnel selection, individualized training, and serve to inform system design as baseline resource data to ensure the system $R_D \leq R_A$.

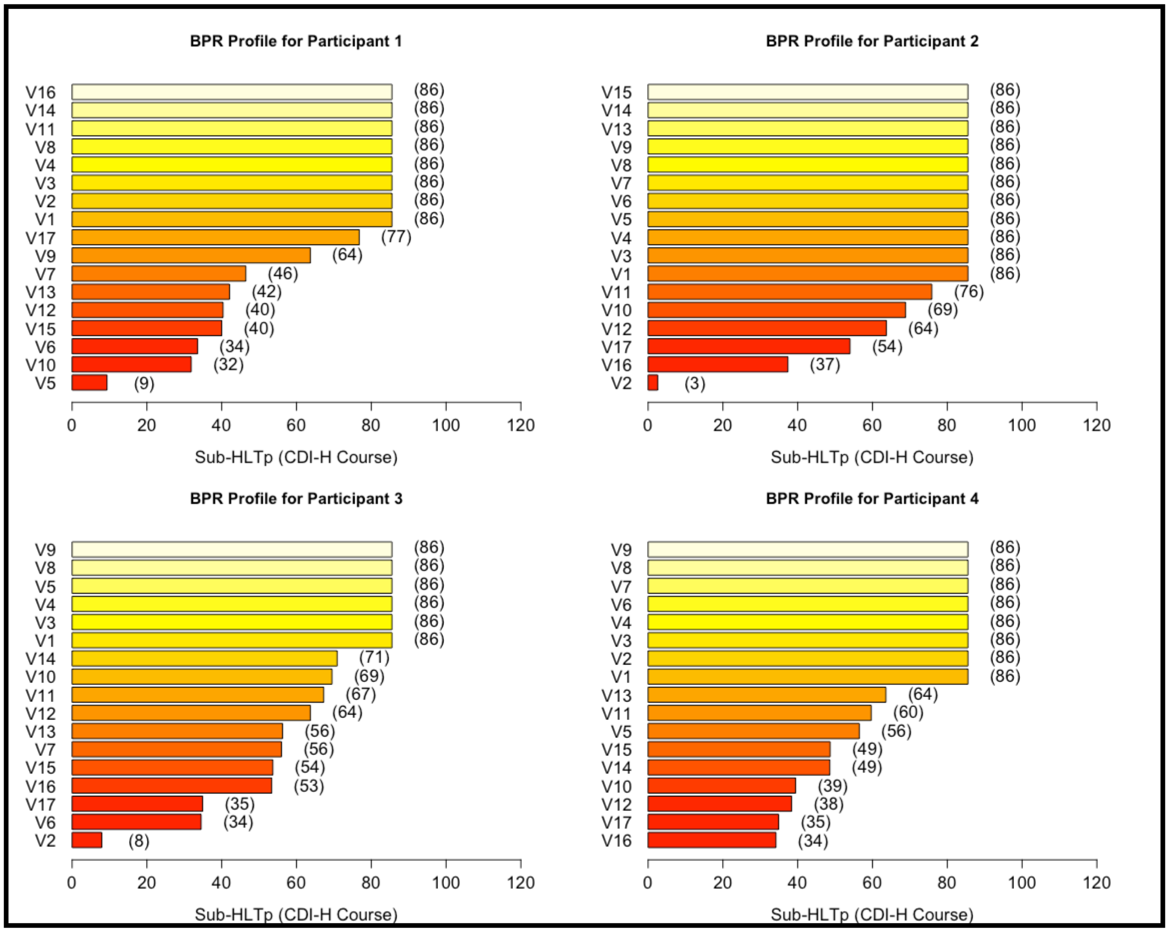


Figure 92. Course (V49) Sub-HLT BPR Profile Sample

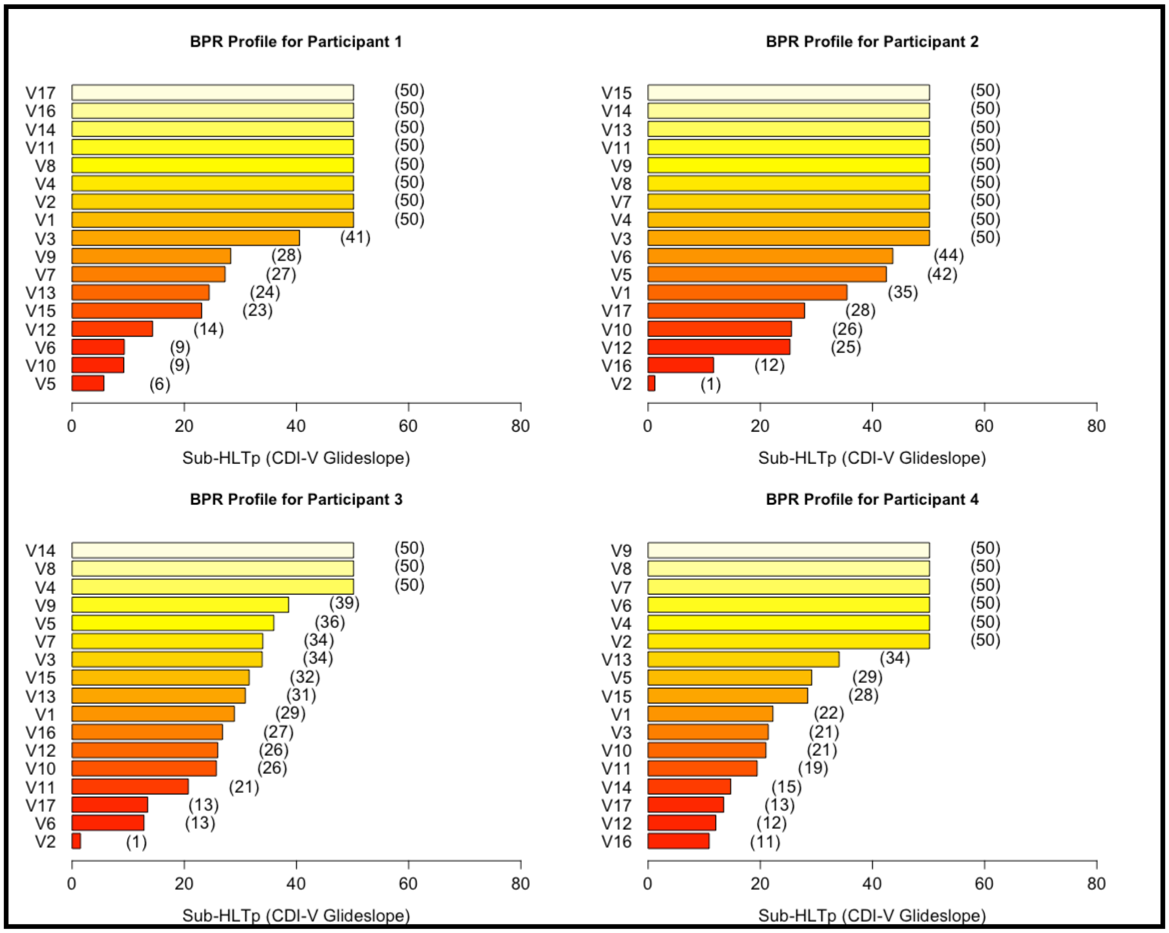


Figure 93. Glideslope (V54) Sub-HLT BPR Profile Sample

b. Discussion

The small C.2 sample set (n = 7) may threaten the validity of the C.2 HLTp 1 *Median* scores. Specifically, three of the larger HLTp 1 scores in the entire study set randomly appear in C.2 (6,064.8, 9,696.37, and 14,015.29). Not surprisingly, comparison between Track A (Control) and C.2 scores demonstrated significant differences according to WRST [$W_{40,\,7} = 2$, $P = 0.000$]. Additional C.1 and C.2 participants would assist in validating these findings. They would also produce updated R_D HLT profiles for the new HLT as modified. These new profiles would identify the new limiting BPRs reflecting this new HLT. These insights further support that stakeholder-determined HLTp is critical during the iterative system design process. The data also suggests this approach may increase the capacity of HSI to enable MBSE in system design, training, and personnel selection in terms of TSP.

In closing, Project III demonstrated the feasibility of using the models built earlier in the dissertation to empirically evaluate HSI domain trades by changing R_A and R_D. These effects adjust constraint boundaries for determining optimal solutions. Perhaps a hybrid approach that seeks to optimize both sides of the resource economic equation is possible. This might include individualized training and targeted design changes in support of TSP. At a minimum, these methods and the results appear to be encouraging in terms of diagnosing system performance issues and delivering targeted remedies to improve TSP and minimize costs. The results of Project III support Chapter VII in terms of facilitating mathematical program formulation.

VII. ADDRESSING A HIGHER STANDARD: OPTIMIZATION

> The final test of a theory is its capacity to solve the problems which originated it.
>
> —Dr. George Dantzig
> The father of optimization theory, 1914 -2005

A. INTRODUCTION

George Dantzig's twin discoveries of linear programming and the simplex algorithm (c. 1947) "have enabled mankind for the first time to structure and solve extremely complex optimal allocation and resource problems" (Dantzig & Thapa, 2006). Dantzig's discoveries, generalized as mathematical programming, provide the capacity to "state general goals and lay out a path of detailed decisions to be taken in order to best achieve these goals when faced with practical situations of great complexity" (Dantzig & Thapa, 2006, p. xxxiii). What spawned from Dantzig's discoveries was the now well-established and proven practice of operations research (OR), which studies "how to form mathematical models of complex engineering and management problems and how to analyze them to gain insight about possible solutions" (Rardin, 2016, p. 1).

In his 2010 dissertation titled *A Discourse in Human Systems Integration*, Tvaryanas states that "technological complexity and its effects on personnel [suggests] the fundamental impetus for HSI was complexity" (Tvaryanas, 2010, p. 530). This highlights the human-machine resource mismatch problem Human Systems Integration [HSI] must address. Therefore, if HSI is to address complex allocation and resource problems, and if OR demonstrates capacity to solve complex problems, then Model-Based Human Systems Integration (MBHSI) cannot avoid OR's proven methods. The Department of Defense's (DOD) stated optimization objective for HSI echoes this requirement. Thus, the purpose of this chapter is to reveal how MBHSI may facilitate mathematical program formulation (optimization) to address total systems performance (TSP) in the DOD. This purpose aims to support the DOD objective of HSI and describes how MBHSI communicates its value proposition in terms of engineering to Model-Based Systems Engineering (MBSE) and its containing systems. Figure 94 once again illustrates the MBHSI system hierarchy, depicting its containing systems and its enabling relationship with MBSE.

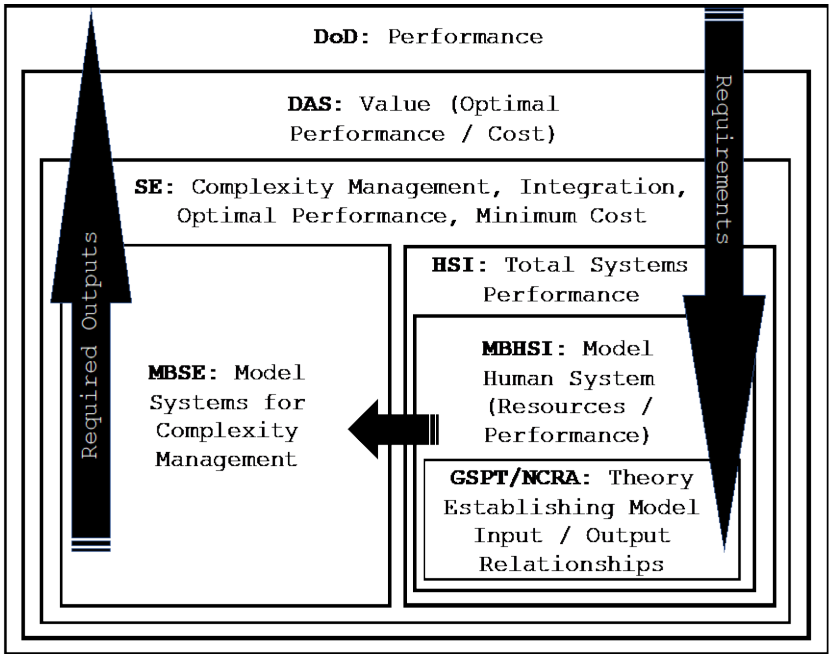


Figure 94. MBHSI Requirements Hierarchy and Related Outputs. Adapted from Hitchens (1992, p. 10).

The goal of this chapter is to describe how MBHSI seeks to inform the three dimensions of mathematical program formulation (decision variables, constraints, and objective functions). These variables are critical because they are the keys that unlock the suite of OR methodologies. In this chapter, an examination of optimization theory partially establishes the relationship between MBHSI and OR. A brief overview of Dantzig's "row" approach to formulation maps MBHSI outputs to this well-established methodology as an example (Dantzig and Thapa (2006). Project III supports system intuition by investigating changes to resource values. These changes demonstrate resource constraint relaxation and restriction to articulate the trade space. Sample outputs from Project II suggest insight gained by MBHSI regarding variables that employ optimization. These accurate and reliable valuations describe how MBHSI facilitates mathematical program formulation. These methods address a higher standard in terms of TSP: optimization. This improved

HSI capacity suggests that this complementary model-based approach to MBSE has merit and deserves further effort to continue its development. Figure 95 provides a conceptual roadmap for this chapter.

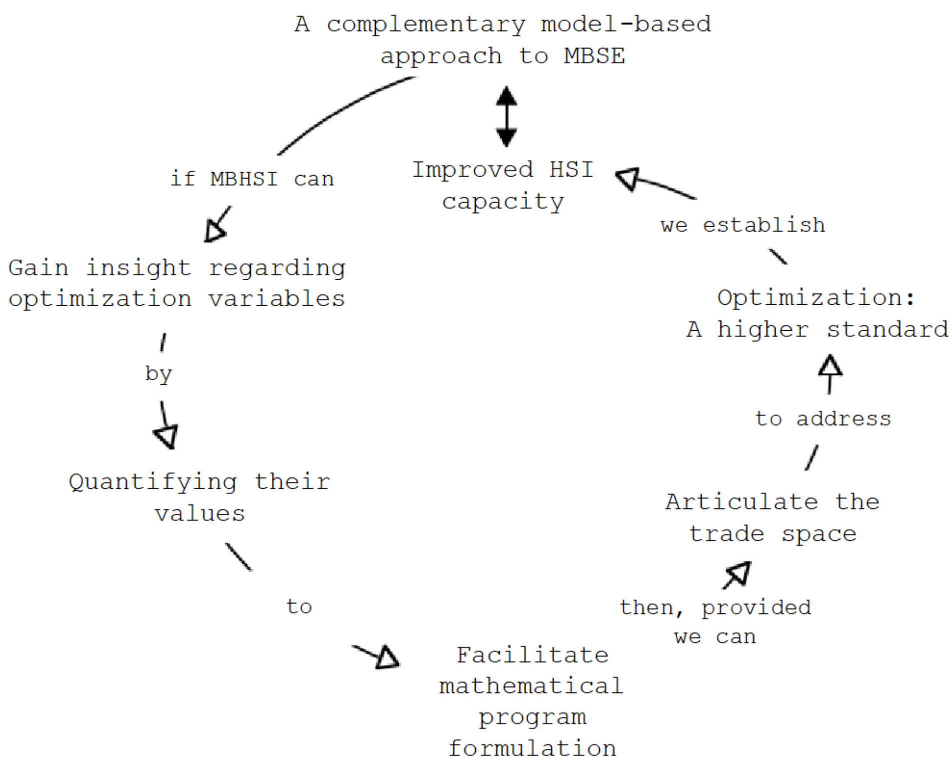


Figure 95. Chapter VII Conceptual Roadmap. Adapted from Hitchens (1992).

B. PROBLEM STATEMENT

DoDI 5000.02PR (draft) requires HSI to “integrate human considerations into system design to optimize total system performance and minimize life-cycle costs” (Department of Defense, 2020, p. 1). HSI in the DOD might benefit from more reliable methods to facilitate mathematical programming (i.e., optimization). Specifically, a method for quantifying optimization resource variables for HSI may reliably facilitate mathematical program formulation. To address this policy requirement and to support HSI containing systems, HSI needs to develop complementary methods for modeling, forecasting, quantifying HSI trade space, and addressing optimization. This chapter offers a theoretically based process for HSI to engage in true optimization. This process of quantifying constraints, informing decision variables, and supporting objective functions opens up MBHSI to the proven OR methods. These insights may increase the capacity for HSI to address the optimization challenges that persist in the DOD in terms of HSI.

C. OBJECTIVE

This chapter articulates the source of variable values in terms of mathematical program formulation using resource and outcome data from Projects I-III. If MBHSI can communicate quantitatively regarding the three fundamental concerns of OR models (decision variables, constraints, and objectives), then truly optimal solutions may be possible in terms of TSP.

D. OPTIMIZATION THEORY AND MBHSI

In establishing a relationship between MBHSI and OR, review of optimization theory in the context of MBHSI sheds light on the importance of identifying values for OR model variables. In *Linear Programming*, George Dantzig defines mathematical programming (or optimization theory) as "that branch of mathematics dealing with techniques for maximizing or minimizing an objective function subject to linear, nonlinear, and integer constraints on the variables" (Dantzig & Thapa, 2006, p. 1). Therefore, if MBHSI is to truly engage in optimization, then constraints on these variables must be accurately and reliably quantified. This dissertation has demonstrated a potentially reliable methodology for addressing this requirement. Specifically, General Systems Performance Theory/Nonlinear Causal Resource Analysis (GSPT/NCRA) establishes linear constraints because it models systems based on capacity. This theoretically based approach might productively engage optimization theory to communicate MBHSI's value proposition in engineering terms.

A closer look at a form of mathematical programing that focuses on linear constrains, linear programming (LP) identifies potential alignment between MBHSI and OR. Dantzig (1997) defines linear programming (LP) as being "concerned with the maximization or minimization of a linear objective function in many variables subject to linear equality and inequality constraints" (p. 1). This form of programming seeks to gain intuition about system behavior and prescribes, according to Dantzig (1997), “actions to be performed by the system so that it may move from its given status towards some defined objective” (p. 1). The two primary objectives for HSI include maximizing performance and minimizing costs. These outcome and cost variables explicitly map to the value equation introduced in Chapter I, where value is the ratio of outcome over cost. Therefore, alignment between OR/LP objective functions and MBHSI can be described as MBHSI providing LP constraint values to facilitate optimization modeling.

A review of the OR process sheds light on the iterative nature of OR modeling and what it seeks to achieve. According to Rardin (2016), “Operations research deals with decision problems [like those engaged by HSI and SE] by formulating and analyzing mathematical models-mathematical representations of pertinent problem features” (p. 3). He also describes the OR process as beginning with program formulation, which includes defining variables and quantifying the “relationships needed to describe relevant system behavior” (p. 3). Next, OR modeling expertise and technology leads to analysis of model outputs where conclusions come from the model, not the problem it represents. Then, inference asks if the “conclusions drawn from the model are meaningful enough to infer decisions for the person or persons with the problem” (p. 4). This process repeats until implementation needs are satisfied. The nature of the iterative process allows for changing variable values, thereby creating an evolving solution space. For example, in Project III, constraint relaxation and restriction efforts resulted in different feasible solution spaces. One might decide, based on Project III evidence, to pursue Glideslope automation. If so, the cyclic modeling process would continue as the model seeks outputs from the adjusted program (i.e., new RDFs) calibrate to the modified HLT R_D. Figure 96 illustrates this OR process.

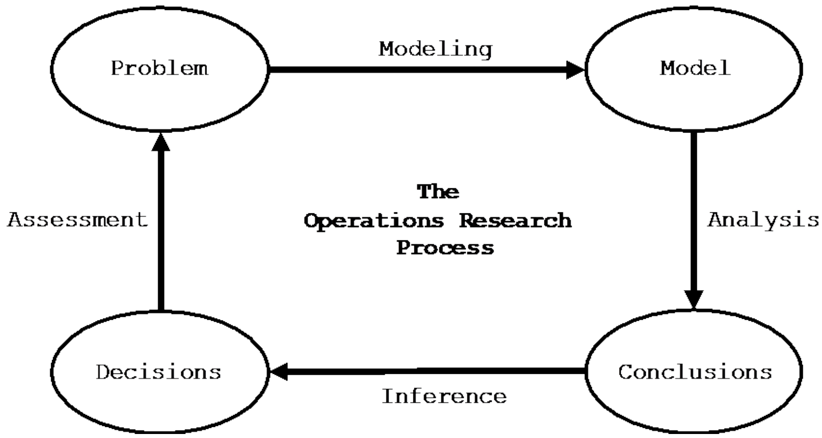


Figure 96. The Operations Research Process. Source: Rardin (2016, p. 3).

The OR approach to decision-making develops mathematical models for real-world problems. However, Dantzig reminds us that "before you can put a problem into a computer and efficiently find a solution, you must first abstract it, which means you have to build a mathematical model" (Dantzig & Thapa, 2006, p. 8). When engaging with nearly any "decision problem - e.g., engineering, business, personal-explicitly defining the decisions, constraints, and objectives helps to clarify the issues" (Rardin, 2016, p. 4). These three fundamental dimensions of mathematical program formulation place abstraction efforts of the problem up front. Decision variables, such as system manpower allocation, design configuration, personnel entrance criteria, training system quantity, or safety features, are open to the decision makers. Constraint variables are those things that limit decisions, while objectives are those things that prioritize decisions. For example, constraint variables may include minimum performance criteria, resource requirements, resource availability, and cost limits, while objectives may be to either maximize performance or minimize cost, or both. This process of abstraction leads to Dantzig's definition of the mathematical model, "the collection of mathematical relationships which, for the purpose of developing a design or plan, characterize the set of feasible solutions of the system" (Dantzig & Thapa, 2006, p. 8).

A key concept to the OR process and systems thinking suggests that adequate investigation surrounding the problem space is a valuable investment in outcomes. Dantzig echoes that the critical process of mathematical model building "is often considered to be as important as solving it because this process provides insight about how the system works and helps organize essential information about it" (Dantzig & Thapa, 2006, p. 8). Rardin (2016) defines optimization models as those that "represent problem choices as decision variables and seek values that maximize or minimize objective functions of the decision variables subject to constraints on variable values expressing the limits on possible decision outcomes" (p. 4). These models "should be such that the decision-making process is not affected by personal bias, whim, emotions, and guesswork" (Balakrishnan et al., 2012, p. 2). Therefore, model variables require accurate valuation.

The quantitative MBHSI approach demonstrates an appropriate methodology to inform decision variables and assign values to constraints in support of the model objective function. In other words, the system PCE and HLT resource demand (R_D) data provide constraint variable values during mathematical program formulation. The system PCE defines the feasible solution space for the HLT and the HLT sets the resource demand signal for the system PCE. This revolution in HSI capacity addresses all three of the optimization problem dimensions during program formulation in support of TSP and cost.

In summary, optimization requirements and MBHSI outputs converge in support of mathematical program formulation. As Dantzig suggests, the mathematical definition of a linear program is to find values of decision variables (i.e., HSI domains) and either maximize TSP or minimize cost while satisfying non-negative resource constraints (R_A and R_D) (Dantzig & Thapa, 2006, p. 7). This dissertation has demonstrated that MBHSI accurately derives resource values that appear to satisfy LP constraint requirements. This facilitates a multitude of mathematical program formulations in terms of HSI decision variables. This may improve HSI's value proposition to MBSE and its containing systems in terms of HSI's mandated objectives, HSI policy, containing systems requirements, and this dissertation's MBHSI-FRs.

E. FORMULATING LINEAR PROGRAMS

Chapter II of relayed that Dantzig (Dantzig & Thapa, 2006) describes two approaches to formulating LPs: column and row. Different system views may lend to preference of one approach over the other. He first describes the column approach and then contrasts it with the row approach. It appears that MBHSI might benefit from using the row approach because of the agreement between limiting resources. In this section, the researcher foreshadows how MBHSI might facilitate “Dantzig’s Row Approach” to LP formulation. Each of Dantzig’s steps are provided and include a short description of how MBHSI might support each step of the row approach:[21]

1. Define the Decision Variables—“Define all the decision variables that represent the quantity to buy, produce, etc.” (p. 11), for example, the manpower quantity to operate a system, the number of aircraft to build, or trained personnel to produce.

    MBHSI informs decision variable definition through identification of HSI domains (i.e., HSI domains critical to address the objective function). The variables are measured by HSI domain resources (R_A and R_D).

2. Define the Item Set—“Determine the classes of objects, the items that are required inputs or are produced as outputs,” units of measure required for each type. Choose only those items that are “bottlenecks” (p. 11). This includes costs.

    MBHSI informs the item set definition through identification of limiting BPR capacities (measured by BPR). Costs (measured in either dollars or R_D).

3. Set Up Constraints and the Objective Function—“Write down the constraints associated with the bottleneck by noting how much of each item is used [R_A] or produced [HLTp] by a unit of each decision variable” (p. 11). For example, decision variables represent different system configurations, how much R_A is used by each configuration or how much HLTp is produced.

[21] This series of steps in formulating mathematical programs strongly influenced the MBHSI-FRs as it details the requirements to achieve optimization. Recall from Chapter I, the MFRs seek to ensure MBHSI satisfies not just containing system requirements, TSP, policy requirements, but also optimization.

MBHSI informs the Input-Output Coefficients through measurement of R_A and R_D as inputs and HLTp as outputs. "Cost leads to the objective function to be minimized" (p.15).

"The capacity items [constraints] each lead to inequality constraints" (p. 15). For example, if four HLT configuration types are being evaluated, each expresses a unique R_D, which according to our resource economic equation ($R_A \geq R_D$) and LP formulation requirements must be less than or equal to R_A. Another example might include personnel alignment in a system where R_D limit entrance into a specific HLT (MOS).

Figure 97 illustrates the row approach to LP formulation. This approach was introduced in Chapter II of this dissertation.

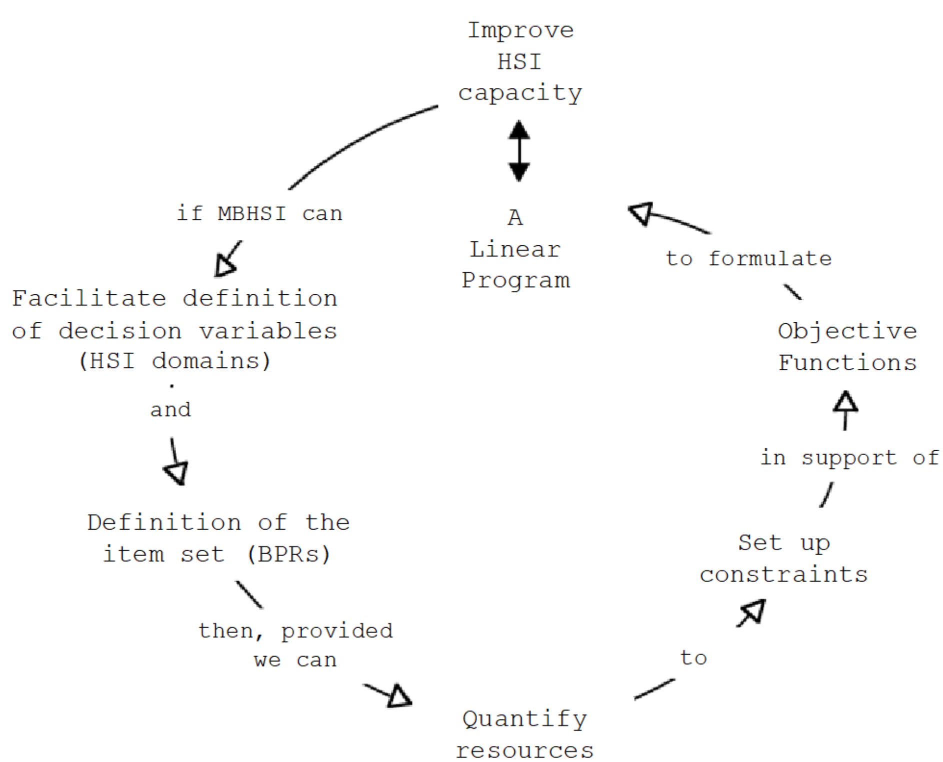


Figure 97. Using MBHSI to Facilitate Mathematical Program Formulation. Adapted from Hitchins (1992).

Figure 98 provides a high-level description of MBHSI mathematical program formulation. The figure illustrates the BPR data collection, HLTp output data collection, R_A and R_D profiles, then seeks cost data to facilitate a program. Obviously, there are many problem sets; this simply describes how MBHSI informs optimization variable values as described.

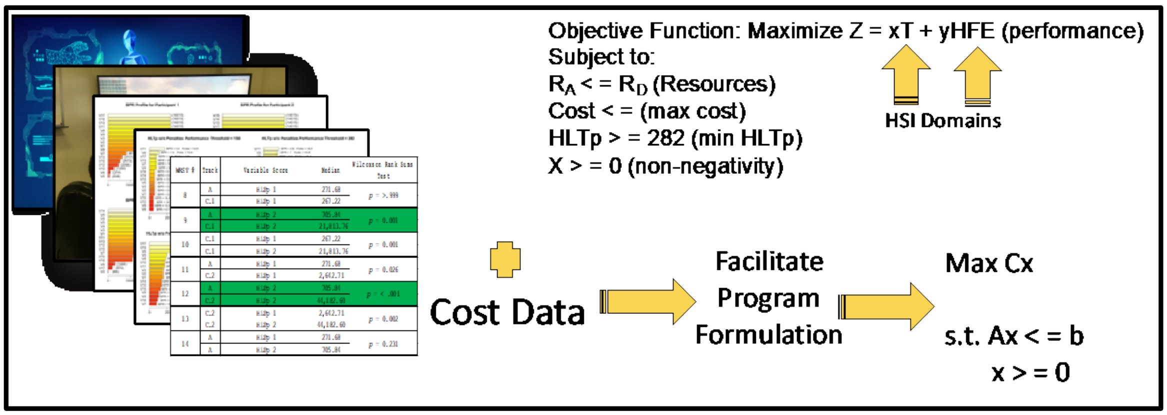


Figure 98. MBHSI Mathematical Program Formulation Concept of Operations

Different variable values result in feasible, infeasible, or even optimal solutions. Feasible solutions represent "a choice of values for the decision variables that satisfies all constraints" (Rardin, 2016, p. 7). Infeasible solutions violate constraint boundary conditions, these are commonly known as "problems." Rardin (2016) defines optimal solutions as "feasible solutions that achieve objective function value(s) as good as those of any other feasible solutions" (p. 7). Unfortunately, many HSI decisions may fall into the infeasible category largely because constraint values may be either vaguely known or perhaps disregarded. Thus, if feasible and/or optimal HSI solutions are required, then variable constraint values should not only be known, but preferably quantified.

Quantifying constraint variables help establish system boundaries, or the defined solution set, which may not always be fixed. Resource values change, new technologies change R_D for a given task, PCEs change over time, and R_A is subject to a multitude of inputs. Rardin (2016) warns us that "if we took nothing as settled, models would mushroom in complexity and meaningful analysis would become impossible" (p. 10). However, he also states that we must "recognize the inherent arbitrariness in system boundaries" (p. 10). Thankfully, OR addresses this issue in one way by engaging in sensitivity analysis. This "exploration of results from mathematical models evaluates how they depend on the values chosen for parameters" (Rardin, 2016, p. 11). This analysis quantifies sensitivity among the parameter-result relationship. Project III examined and articulated effects of parameter changes on both sides of the resource equation, R_A and R_D. These changes to a LP demonstrate an approach to seeking insight into a system.

F. COMMUNICATING THE MBHSI VALUE PROPOSITION

Figure 99 illustrates an example of MBHSI resource constraint values (R_A) with precision at the individual-level. These values facilitate definition of the feasible solution space in terms of LP. Specifically, the identified limiting BPR defines a constraint for each individual. For Participant 1, V5 is the limiting BPR while Participant 2 is limited by BPR V2. While many constraints do not demonstrate binding (bottleneck) characteristics, those resources that do, define the feasible solution space. This reality explicitly demonstrates the value of GSPT/NCRA as a unifying construct for MBHSI. Optimization's dependence on constraint variables (system boundaries) and GSPT's focus on limiting BPRs (system capacities) explain the potentially powerful relationship between OR and MBHSI via GSPT. Additionally, Figure 100 illustrates a sample of resource constraint values (R_D) at various HLTp levels. These values facilitate LP formulation by quantifying resource requirements given a desired HLTp. Specifically, the values derived detail the amount of a resource required to accomplish the HLT to a specified level of HLTp. For an HLTp of 150 or 282, the BPR with the highest demand is V11 but at an HLTp of 1500 or 5000, the BPR with the highest demand is V17. The relationships between each constraint (R_A and R_D) informs multiple types of OR decision problems. For example, a personnel alignment OR model might seek to minimize training costs or maximize HLTp by optimizing R_A and R_D allocations. Perhaps an individual training OR model might seek to develop optimal individual training plans that target specific HLTp levels in an effort to minimize training costs and maximize TSP.

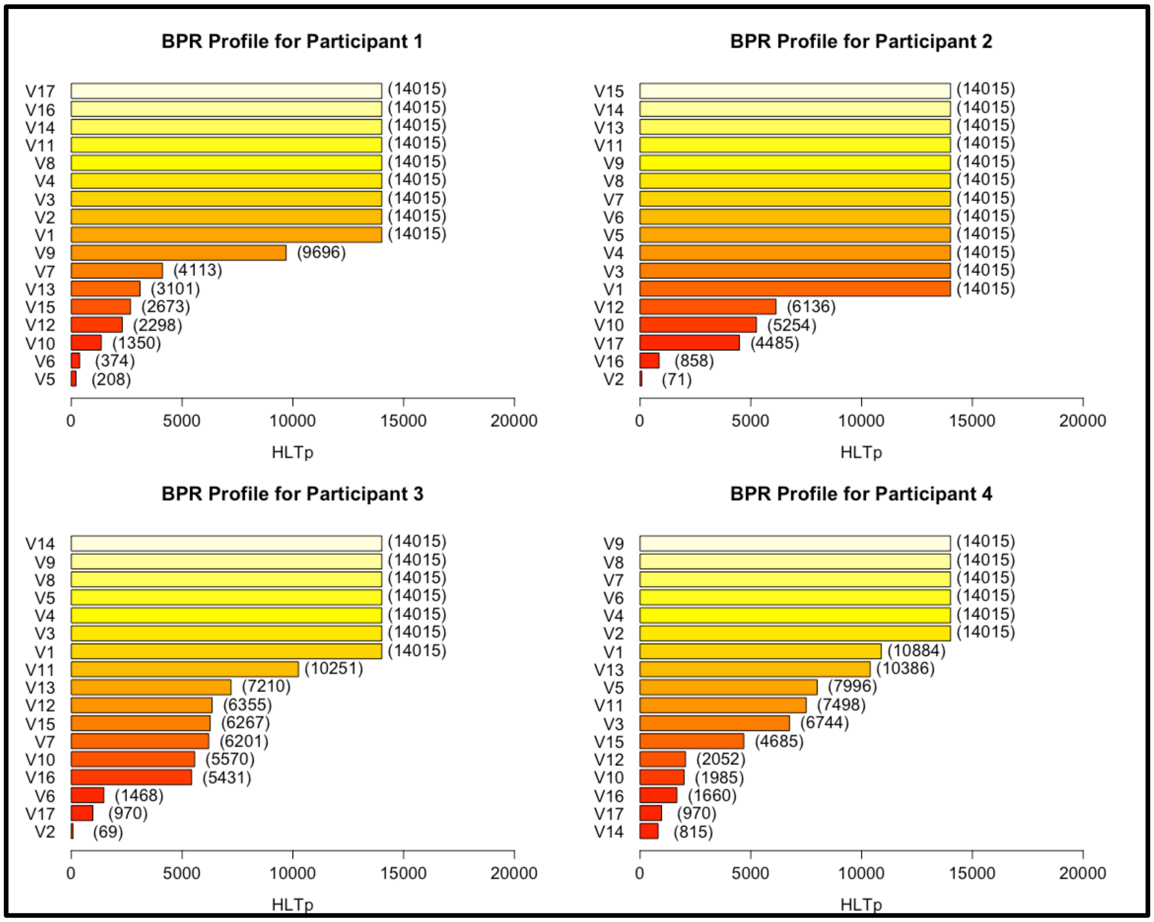


Figure 99. Sample of LP Constraint Variable Values (R_A)

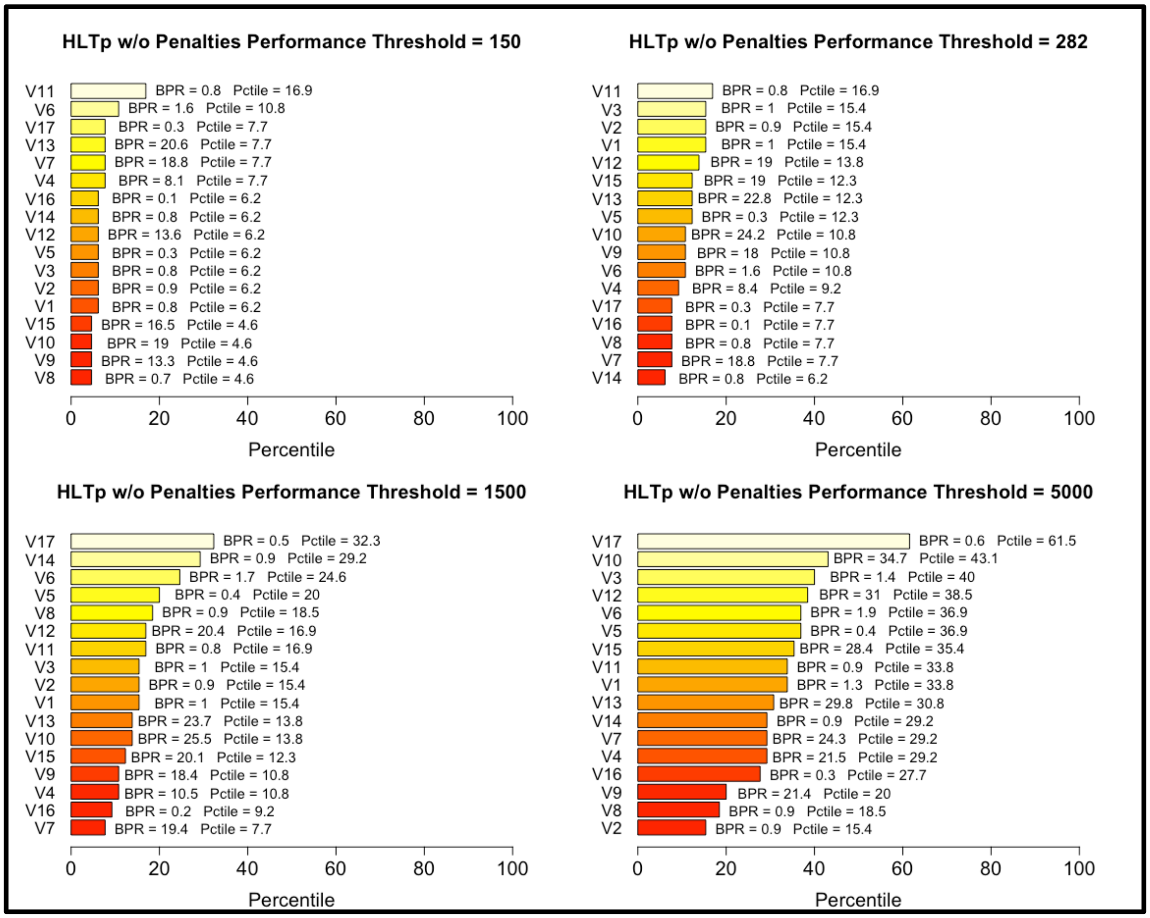


Figure 100. Sample of LP Constraint Values (R_D)

Project II explored MBHSI's capacity to quantify relationships between resources and performance to characterize the feasible solution set. Forecast accuracy supports study model validation. Figure 101 illustrates the RDF concept using an individual performance forecast given a single BPR from the study. Specifically, the green vertical line represents observed HLTp, the nonlinear blue line represents the MBHSI RDF, and the red line predicts HLTp given BPR value. In this particular case, MBHSI slightly over-predicts HLTp. This form of "abstraction" by MBHSI facilitates the building of an OR mathematical model by explicitly and reliably defining constraints.

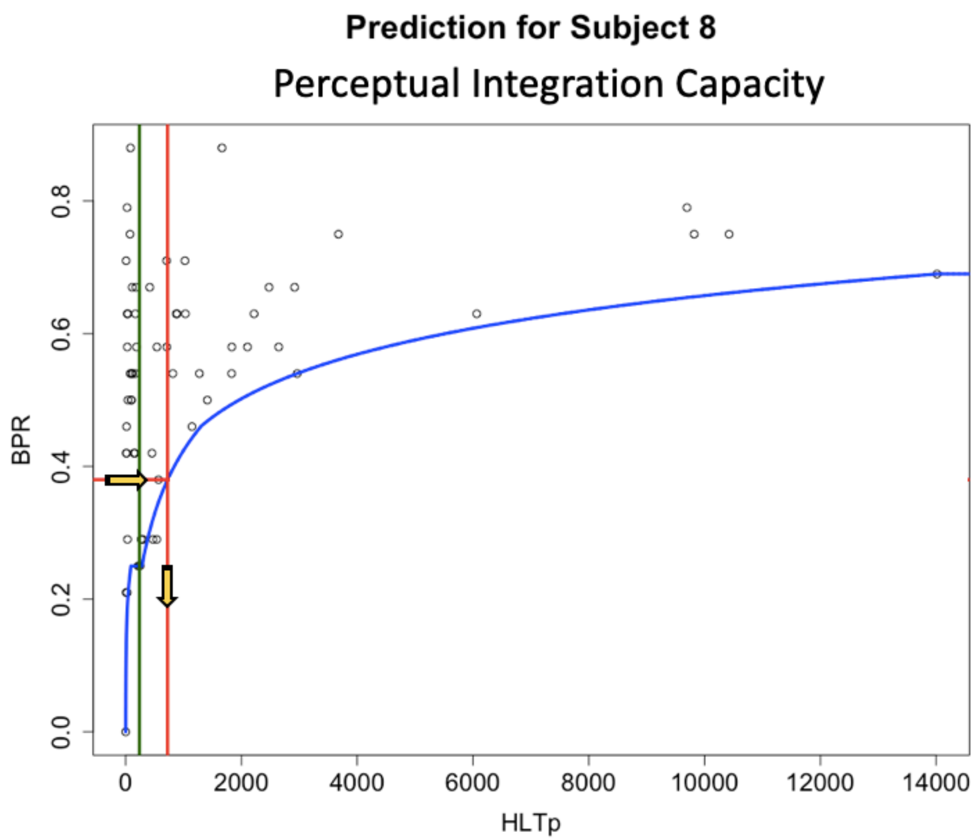


Figure 101. Individual Performance Forecast Illustrating MBHSI Abstraction

In closing, Dell (2018) has said that "all real-world problems are infeasible." This reality warrants careful study regarding HSI and its support to SE. HSI in the DOD has a clear impact on outcomes and cost: it influences value. Thus, addressing the higher standard of optimization, while difficult, may prove to add value as proffered by Rardin (2016): "Operations research is founded on the conviction, buttressed by a long history of successful practice, that formulation and analysis of mathematical decision models is often worth the trouble" (p. 20).

VIII. DISCUSSION

> If we don't communicate with the outside world to gain information for knowledge and understanding we die out to becoming a non-discerning and uninteresting part of that world.
>
> —Col. John Boyd (1927–1997)

A. ORIENTATION

A review of the dissertation thesis statement along with the Model-Based Human Systems Integration (MBHSI) mission statement and definition highlight the contributions of this work. The appraisal of this new Human Systems Integration (HIS) concept engages the MBHSI functional requirement (MBHSI-FR). This appraisal is important to "check for reversibility and match-up with reality [to] demonstrate internal consistency" (Boyd, 1976, p. 3). The MBHSI approach is described as containing system requirements. A review of overarching conclusions provides key take-aways, lessons learned, threats to validity, and limitations. A section titled "Next Steps" paves a way forward for MBHSI in the Department of Defense (DOD). Identification of potential future work provides a map for further exploration. Figure 102 illustrates the chapter roadmap.

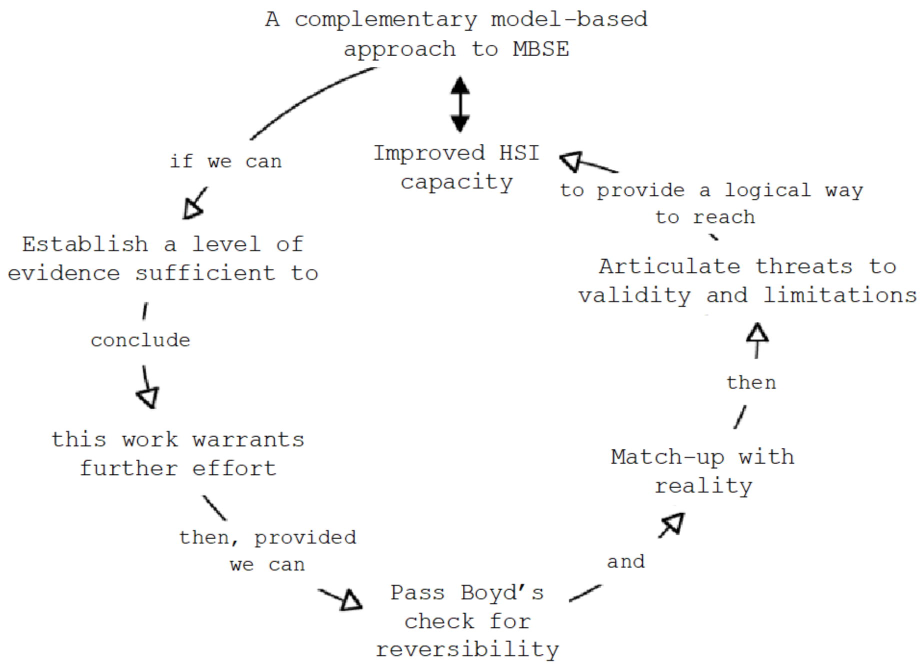


Figure 102. Chapter VIII Roadmap. Adapted from Hitchens (1992).

B. CONTRIBUTIONS

First, the reader deserves an explanation as to why I undertook this work. Col. Boyd's *Roll Call*, "To Be or To Do," describes a fork in the road where we must make a decision about which direction to take. "If you go that way you can be somebody, or you can go that way and you can do something-something for your country and for your Air Force and for yourself" (Hammond, 2012, p. 12). At the onset of my doctoral studies, I encountered Boyd's *Roll Call*: I had a choice to make. I emphatically and unapologetically chose "To Do," despite the fact there were easier paths to the destination: graduation. I am driven by the desire to change people's fundamental understanding of HSI, and sincerely want to make a significant, long-lasting contribution to the warfighter, fellow Airmen, and my United States Air Force through the practice I have been fortunate to be a part of: Human Systems Integration.

Recall the dissertation thesis statement from Chapter I.

> HSI lacks a generally accepted unifying theoretical perspective that joins HSI domains in terms of overall systems performance. The warfighter, HSI,

> SE, and the DOD would benefit from a theoretical perspective that bridges domain considerations with system performance in terms of HSI. General Systems Performance Theory (GSPT), defined as "a framework for modeling systems, tasks, and their interface using an abstraction that focuses on performance and all attributes thereof," is proposed as a unifying theoretical perspective for HSI *and* MBHSI. (Kondraske, 2011, p. 238)

This dissertation engaged Boyd's (1976) "unstructuring and restructuring" to develop a new concept of HSI. An improved and standardized approach to datafy (model) the human system in terms of capacity to accurately forecast performance, articulate the HSI trade space, and facilitate optimization program formulation was explored. This improved approach, evidenced by the results of the three iteratively built projects, suggests an increase in HSI's capacity to deliver order and meaning for systems engineering (SE) and the DOD may be possible. The guiding mission statement was

> To create an orderly, sensible, theoretically and model-based methodology to develop truly integrated solutions for the warfighter at minimum cost, optimizing the conversion of resources into TSP.

The modularized process of this dissertation created this methodology. The results suggest promise in facilitating mathematical programming and supporting optimal solutions for the warfighter and the DOD. Unstructuring optimization led to the identification of specific requirements, including the valuation of optimization variables. This valuation defines the outputs of HSI to OR modeling. This requirement was targeted through the engagement of General Systems Performance Theory/Nonlinear Causal Resource Analysis (GSPT/NCRA) to explicitly establish the relationships between HSI domain concerns re-expressed as resource issues, and total systems performance (TSP). The quantitative values that generate insight and understanding regarding these relationships appear to cleanly address this critical requirement. Recall that operations research (OR) is defined as "the study of how to form mathematical models of complex engineering and management problems and how to analyze them to gain insight about possible solutions" (Rardin, 2016, p. 1). Therefore, OR appears to be the link between the gulf Hitchens (1992) describes between the human factors community (including HSI) and SE. This is important because the outputs of OR mathematical models communicate clearly with MBSE and SE. MBHSI and OR model outputs are how MBHSI communicates its value proposition to MBSE. Figure 103 illustrates these critical relationships.

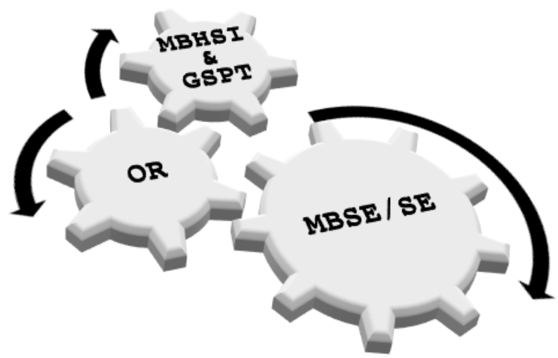


Figure 103. Operations Research Relationship with MBHSI and MBSE

The result of this creative induction, MBHSI, was defined in Chapter III:

> MBHSI is an essential, model-based, and integrative process that reliably addresses complexity in terms of resource economics while enabling the SE practice. It applies GSPT and NCRA to model and forecast the quantitative relationships between HSI domain resources and system-level performance, targeting the chronic HSI trade space problem and the original objective of HSI, optimization. Finally, it seeks to communicate its engineering and program management value in engineering terms.

C. MBHSI FUNCTIONAL REQUIREMENTS

Boyd (1976) suggests that if humans will “agree to constraints in order to collectively pool skills and talents, obstacles can either be overcome or removed” (p. 1). If a practice is unable or unwilling to identify a robust underpinning (unifying theoretical perspective) that conforms to the containing system, obstacles standing in the way of HSI goals may become permanent and continued marginalization may be likely (Booher, 2003; Hitchins, 1992; Tvaryanas, 2010). Boyd suggests that “alienated members may dissolve their relationship and form a new group to improve capacity for action” (p. 1). Deducing HSI’s mandated objective—optimization—through the lens of HSI’s purpose as stated in DOD policy and containing system requirements led to the development of the MBHSI-

FRs. Satisfying these MBHSI-FRs defines how MBHSI informs TSP. An appraisal of MBHSI against these functional requirements serves as an appropriate check for reversibility and to demonstrate internal consistency (Boyd, 1976). Figure 104 provides each of the MBHSI-FRs mapped to the laboratory study projects.

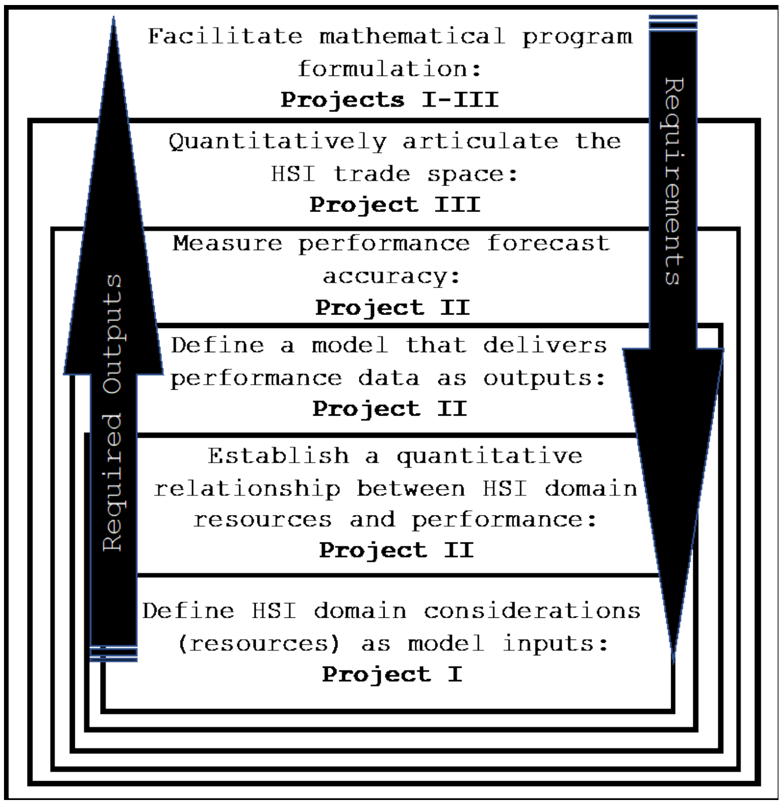


Figure 104. MBHSI-FRs Mapped to Laboratory Study Projects

(1) MBHSI-FR #1: Define HSI domain considerations (resources) as model inputs

Project I engaged GSPT to define HSI resources as model inputs. Model inputs are defined as BPRs. GSPT measures these inputs in terms of capacity. These terms demonstrate linear constraints on the mathematical program. Therefore, the results from Project I satisfy this requirement.

(2) MBHSI-FR #2: Establish a quantitative relationship between HSI domain resources and performance

Project II engaged the concepts of NCRA and the novel resource demand function (RDF) code developed by committee member Dr. Koyak and the researcher. These RDFs established quantitative and accurate relationships between HSI resources (basic performance resources [BPR]) and performance (high-level task performance [HLTp]). Additionally, the models were validated using a jackknifing technique (leave one out-fit-replace-repeat) for each model across all 17 study BPR measurements. Therefore, the results from Project II satisfy this requirement.

(3) MBHSI-FR #3: Define a model that delivers performance data as outputs

The GSPT/NCRA conditional expectations not only deliver performance data as outputs, a reverse read reliably delivers resource data as outputs. These conditional expectations are expressed as:

$$E\ [HLT_P|R_A].$$

$$E\ [R_A|HLT_P]$$

Therefore, the selection of GSPT/NCRA as a contained system for MBHSI satisfies this requirement.

(4) MBHSI-FR #4: Measure performance accuracy

Project II measured performance accuracy at the individual level and sample set level. The novel code developed for MBHSI included forecast performance accuracy functionality. Specifically, MBHSI measures forecasted performance and compares it to observed performance. Therefore, the results from Project II satisfy this requirement.

(5) MBHSI-FR #5: Quantitatively articulate the HSI trade space

Project III engaged in feasible space relaxation and restriction across two of the study tracks (B and C). In Track B, the study investigated training (IV) and its effect on HLTp (DV) as well as training (IV) and its effect on test scores (DV). In Track C, the study investigated HFE (IV) and its effect on HLTp (DV). Both study sets reflect data quantitatively articulating the HSI trade space. Therefore, the results from Project III satisfy this requirement.

(6) MBHSI-FR #6: Facilitate mathematical program formulation

Chapter VII explored the three primary dimensions of mathematical programming, optimization, and specifically LP. These dimensions include decision variables, constraints, and objective functions. Objective functions are facilitated by alignment with HSI stated objectives to maximize TSP and minimize cost. Decision variables can be considered HSI domains, facilitating formulation. MBHSI resource (BPR) and performance (HLTp) data facilitate the assignment of values for these critical values. Finally, constraint variables are defined by BPR capacities and HLT resource demands. R_A and R_D are each quantified by MBHSI for the purpose of assigning values to these critical variables. Therefore, Projects I-III outputs appear to facilitate mathematical program formulation as detailed in Chapter VII. However, no actual mathematical program formulations are included in this study. Exploration of MBHSI/OR modeling is clearly the next step for MBHSI.

This study addressed each of the MBHSI-FRs. The systems thinking concept of contained systems illustrates how MBHSI might deliver requisite outputs to each of its containing systems. Chapter II deduced policy, containing systems, and GSPT/NCRA in preparation for this creative induction I call MBHSI. The familiar figure (Figure 105) provides one final illustration of this acknowledgement. The key outputs that MBHSI passes to MBSE and SE include quantitative insights regarding resources at play for a given system. Specifically, the identification of what resources may limit individuals from higher HLTp and what resource demands from the system design tax the human system. Additionally, the capacity to gain understanding about the nonlinear relationship between HSI resources and HLTp may provide improved insight for PMs tasked with system design decisions. Also, the capacity to evaluate various design choices against personnel R_A may be valuable outputs of MBHSI. Finally, in terms of threshold and objective values that lead to TSP and cost, MBHSI brings quantitative understanding to how much is required to achieve stated HLTp levels and identifies which R_A or R_D might be a focus for training, personnel selection, manpower, or HFE to minimize cost and maximize TSP. These outputs articulate how MBHSI might improve the value function defined in Chapter I by improving outcomes and reducing cost.

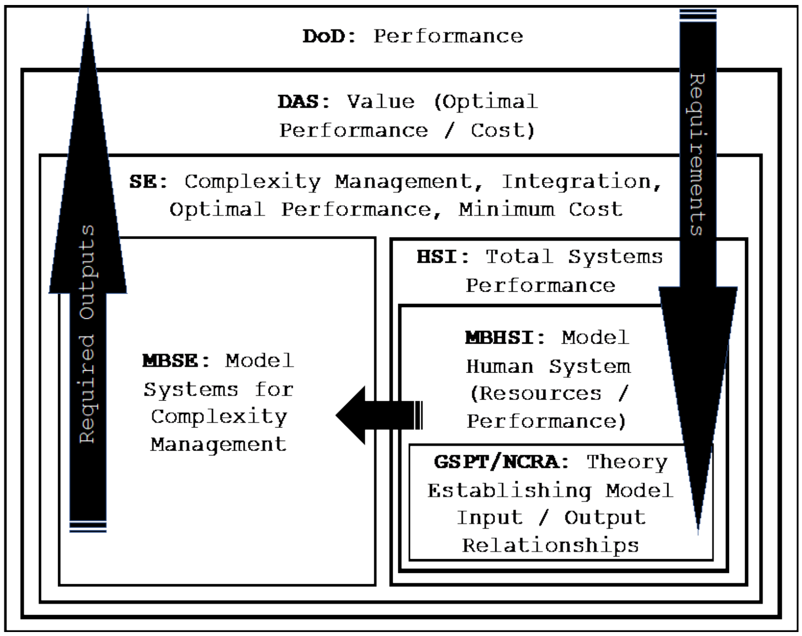


Figure 105. MBHSI Requirements Hierarchy and Related Outputs. Adapted from Hitchens (1992, p. 10).

In summary, this check for reversibility demonstrates internal consistency. The strong appraisal of MBHSI against these functional requirements, based on the outcomes of this laboratory study suggests that MBHSI demonstrates promise as a complementary model-based approach to MBSE. Figure 106 illustrates the foundational nature of the study to engage a unifying theoretical perspective that included three interactively built study projects that must satisfy the MBHSI-FRs to achieve success as MBHSI.

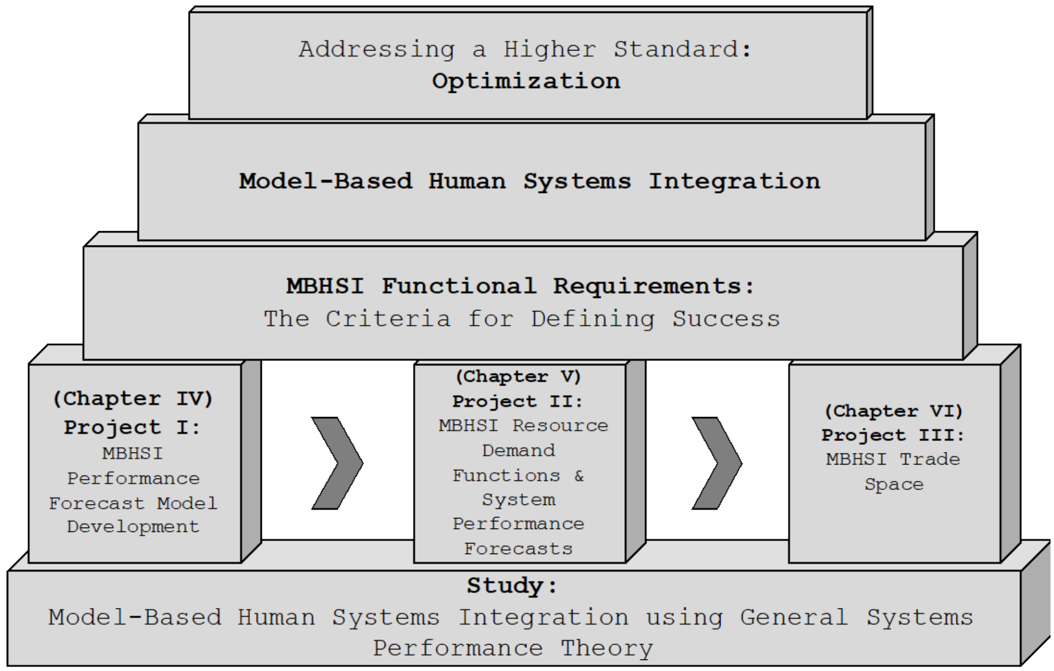


Figure 106. Construction of MBHSI

D. CONCLUSIONS

MBHSI demonstrates the capacity to establish a relationship between the ends (TSP) and means (resources) for HSI. This methodology provides a strong candidate for addressing some of the challenges that persist in the DOD regarding HSI. Furthermore, this study and its results establish potential for a strong relationship with OR to improve communication with SE. The effects-based targeting approach to this dissertation was provided in Chapter I and again here in Figure 107. The advance in HSI capacity as an enabling system to SE is expected to deliver progressive insights to the DAS. This insight enhances what is delivered to the warfighter to improve outcomes and costs. This improvement in outcomes and cost target influence the operational arena.

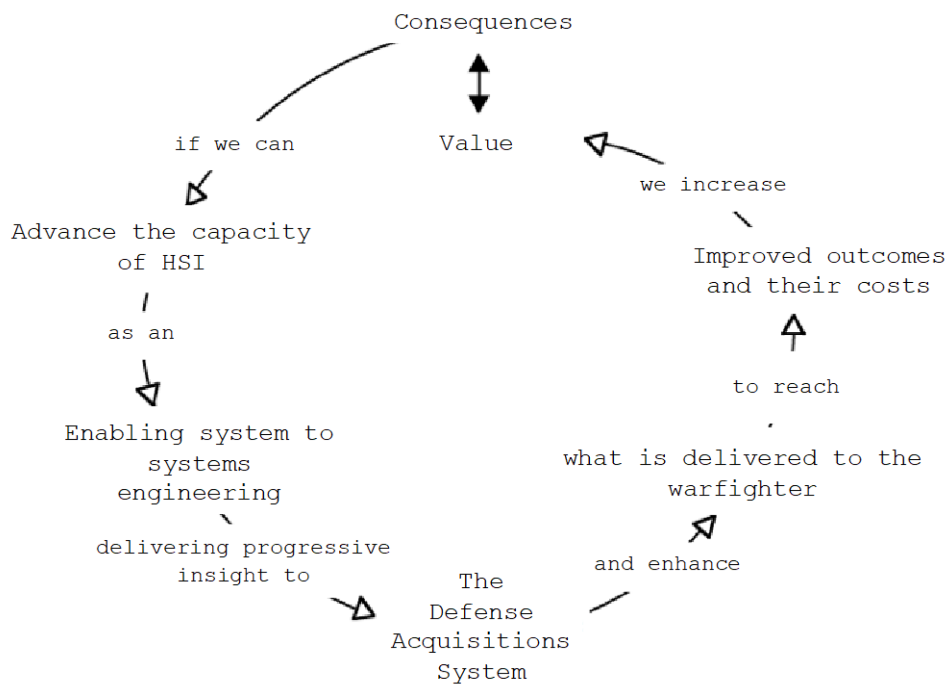


Figure 107. An HSI Effects-Based Approach to Improve System Value in the DOD. Adapted from Hitchens (1992).

E. THREATS TO VALIDITY AND LIMITATIONS

Each project detailed threats to validity and limitations of this dissertation. However, MBHSI is a system and, like any other system, has limiting BPRs. MBHSI shares the same threats as MBSE and the Modeling & Simulation (M&S) communities. Specifically, when modeling a system, an incomplete representation is evaluated, limiting valid outputs. Ecological relationships present in the real-world may not be included in modeling efforts which may threaten model outputs. Additionally, model inputs drive outputs thus, the quality of model inputs must be recognized as a threat to validity regarding outputs. During preparatory phases of this effort, Project I was anticipated to be a limiting BPR, that model inputs would threaten forecast accuracy. This was an accurate assessment in hindsight. That being said, each phase of MBHSI risks being a limiting BPR. In terms of simulation, assumptions made during development limit and/or threaten the reliability of outputs. This was a laboratory study, therefore, its application to a real-world ILS approach should be evaluated in the operational environment. Additionally, MBHSI is currently limited as this dissertation is the first of its kind. Additional work must be accomplished to determine if validity persists. Finally, relentless attention to detail is necessary. I will echo Rardin's OR perspective that "the OR [MBHSI] approach to problem solving works best on problems important enough to warrant the time and resources for a careful study" (Rardin, 2016, p. 20).

F. NEXT STEPS

1. Operationalize MBHSI with Operations Research

The critical next step for MBHSI is engagement with OR experts to explore a spectrum of HSI decision problems and formulate them. This is the operationalization of MBHSI. This dissertation demonstrated that obtaining reliable OR model variable values is possible. Exploration of LP, nonlinear programming, multiple objective functions, minimum cost network flow, efficient frontiers, and goal programming represent but a small sample of proven OR methodologies that may be possible with MBHSI.

2. An MBHSI Practitioner Competence and Body of Knowledge Development Strategy

The practice of HSI has invested significant resources over the years to derive "HSI Competencies" and a "Body of Knowledge." This thrust effort is critical to the practice's evolution in the short- and long-term. MBHSI practitioner competence is a multidimensional outcome that considers practitioner selection, training success, performance in the operational environment, and retention within the practice beyond an initial assignment. The strategy for MBHSI competency development is to simply engage MBHSI on MBHSI. Specifically, define the HLT, HLTp, then measure resources and performance. The following is offered as a way forward.

3. MBHSI Competency Development

This is accomplished through the following steps:

- Define the System (MBHSI practitioner) HLT.

- Establish the minimally acceptable degree(s) of HLTp.[22]

- Identify a sample of practitioners.

- For each individual, measure "*N*" BPRs-i.e., BEPs and/or intermediate level performance resources-in a manner that determines maximum resource availability.

- For each individual, measure HLTp.

- For each BPR, aggregate data for the entire sample and plot observed HLTp vs. measured BPR value and then fit the lower bounds of the data to create an RDF.

- For each BPR, use HLTp and the respective RDF to determine the resource profiles as demonstrated in Project II of this dissertation.

4. MBHSI Body of Knowledge

This concept of a MBHSI Body of Knowledge (BoK) depends on MBHSI performance outcome data. This is analogous to evidence-based medicine. MBHSI must capture how a system design was influenced and how it affected TSP. This collection seeks to capture what was forecasted and the observed outcomes. Additionally, a root cause analysis (RCA) that details why design influences succeeded or failed is critical to any BoK. These evidence-based data become diagnostic tools for future system BPR selection protocols, task analyses in the form of RDFs,[23] and system performance forecast accuracy scores. The BoK is envisioned to be a learning system that matures to relentlessly and quantitatively drive down MBHSI performance forecast Type I and II errors. The BoK should improve practitioner performance standards while also continuously improving the

[22] For the purposes of MBHSI practitioner competencies, the operational community should drive the definition of acceptable HLTp.

[23] A point of contrast with MBHSI is that task analyses are also outputs of the method, not just inputs. Whereas, "major tasks identified in the job task analysis" are considered limiting resources at a given level of performance as outputs as well as inputs when selecting BPRs for measurement.

MBHSI methodology. Defining the appropriate practitioner HLTp identifies the initial set of BPRs. Limiting BPRs inform the MBHSI ISD process.

5. A New Model-Based Human Systems Integration Plan

The new DOD Adaptive Acquisition Framework (AAF) introduced in Chapter II will require an approach to facilitate the six new acquisition tracks. In fact, Section 1.2, "Policy," states that "It is DOD policy that *HSI* [emphasis added] in defense acquisition *provide a disciplined, unified, and interactive approach* [emphasis added] to integrate human consideration into system design to optimize total system performance and minimize life-cycle costs" (Department of Defense, 2020, p. 1). Because MBHSI and GSPT are system agnostic, MBHSI may facilitate the new AAF across the spectrum of acquisition tracks. A robust MBHSI Plan (MBHSI-P) will be necessary going forward. An introductory MBHSI-P is provided:

- Define the System HLT.
- Establish the minimally acceptable degree(s) of HLTp.[24]
- Identify a sample of current or potential operators.
- For each individual, measure "*N*" BPRs-i.e., BEPs and/or intermediate level performance resources-in a manner that determines maximum resource availability.
- For each individual, measure HLTp.
- For each BPR, aggregate data for the entire sample and plot observed HLTp vs. measured BPR value and then fit the lower bounds of the data to create an RDF.
- Forecast HLTp for the defined HLT.

[24] The operational community should support the definition of acceptable HLTp.

- For each BPR, use HLTp and the respective RDF to determine the resource profiles as demonstrated in Project II of this dissertation.

- Articulate the respective trade space across resources and system-level performance.

- Manipulate system BPRs and/or HLT to demonstrate relaxation and restriction of the feasible solution set to inform the next round of this process. This will likely be an iterative process to accommodate changing stakeholder priorities and requirements.

- Engage OR expertise to formulate mathematical optimization models as appropriate.

G. THE LAST WORD

If we are to improve the arena of consequence with HSI, it appears clear we must increase our capacity to communicate more clearly with our engineering partners through established operations research methodologies.

> When you can measure what you are speaking about, and express it in numbers, you know something about it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts advanced the stage of science.
>
> —Lord Kelvin, Irish mathematical physicist, 1824–1907

APPENDIX A. CONSENT FORM

Model-Based Human Systems Integration

Engineering Worthy Performance

Consent Form

August 2019

Naval Postgraduate School
Consent to Participate in Research

Introduction. You are invited to participate in a study entitled *System Performance Forecasting for Model-Based Human Systems Integration*. The purpose of this research is to investigate the accuracy of system performance forecasting and quantitative trade analytics for Model-Based Human Systems Integration using General Systems Performance Theory and Nonlinear Causal Resource Analysis.

1) Participation is voluntary. Refusal to participate will involve no penalty or loss of benefits to which you would otherwise be entitled, and you may discontinue participation at any time without penalty or loss of benefits to which you otherwise would be entitled.
2) Due to the immersive nature of the simulator, slight simulator sickness may be experienced. You will be asked before, during and after the experiment if you feel ill in any way. Potential illnesses you may experience include nausea and dizziness. If at any time you experience any of these symptoms, you will notify the experimenter immediately. In addition to the risks mentioned above, this study will include no more risks than those associated with using a computer or playing a computer-based video game.
3) You will receive no benefits for participating in this study other than gaining knowledge concerning piloting an aircraft. The benefits of the study are expected to inform improved system performance accuracy for system design, development, and operations.
4) You understand that your participation is strictly voluntary, and if you agree to participate, you are free to withdraw at any time without prejudice. If you are uncomfortable participating at any time during the experiment you may withdraw without any retribution. The alternative procedure is to not participate.
5) Participants will be asked to complete a small battery of cognitive and psychomotor tests and to perform real flying tasks in a virtual flight simulator. Your participation in this experiment will last approximately one hour. The experiment will consist of verbal instructions, practical instructions, training sessions, and one to three simulated precision approaches using X-Plane 11 Flight Simulator. You understand that prior to your completion of the High-Level Task, you will complete six cognitive/ psychomotor tests in which your maximum performance is requested. These tests will measure your Basic Performance Resources that are expected to limit HLT performance. The data collected from these tests will not be analyzed for any purpose other than this study, and any findings in this study can in no way disqualify you for flight or military service.

6) Participants will be randomly divided into two within subjects Groups; A and B. Group A participants will build the forecast models. They will also conduct the HLT a second time to generate learning effect data. Group B will be equally divided into Groups B.1. and B.2. Group B.1. will have BPRs measured, HLTp forecasted, HLTp observed, will receive professional training, and complete the HLT a second time to observe a trained HLTp score. Group B.2. will have BPRs measured, HLTp forecasted, HLTp observed, and will complete the HLT a second time with some performance metric automated (Airspeed, altitude, or CDI) to observe a modified HLTp score. The experiment is expected to be conducted with one participant at a time. The participant sample size is expected to be approximately 100.
7) Basic Performance Resource data and simulated flight performance data will be recorded.
8) The experiment will take place in the MOVES Institute Laboratory.

Cost
There is no cost to participate in this research study.

Compensation for Participation
You understand that no tangible compensation will be given. You understand that a copy of the research results will be available at the conclusion of the study by contacting the Primary Investigator Dr. Lawrence Shattuck (lgshattu@nps.edu, 831–656-2473) or the experimenter Matthew Taranto (mttarant2@nps.edu, 702–540-2455).

Confidentiality & Privacy Act
Any information that is obtained during this study will be kept confidential to the full extent permitted by law. All efforts, within reason, will be made to keep your personal information in your research record confidential but total confidentiality cannot be guaranteed. No information will be publicly accessible which could identify you as a participant. You will be identified only as a code number on all research forms/data bases. Your name on any signed document will not be paired with your code number in order to protect your identity. You understand that records of your participation will be maintained by NPS for 10 years, after which they will be forwarded to a federal records center. However, it is possible that the researcher may be required to divulge information obtained in the Course of this research to the subject's chain of command or other legal body.

If you consent to be identified by name in this study, any reference to or quote by you will be published in the final research finding only after your review and approval. If you do not agree, then you will be identified broadly by discipline and/or rank, (for example, "fire chief").

☐ I consent to be identified by name in this research study.

☐ I do not consent to be identified by name in this research study.

Points of Contact. You understand that if you have any questions or comments regarding this
project upon the completion of your participation, you should contact the Primary Investigator,
Dr. Lawrence Shattuck, 831–656-2473, lgshattu@nps.edu. Any other questions or concerns may be addressed to the IRB Vice-Chair, Bryan Hudgens, 656–2043, bjhudgen@nps.edu.

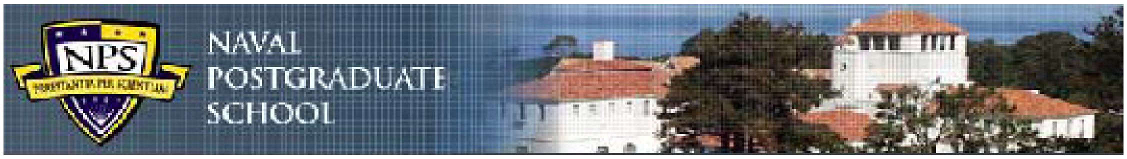
Statement of Consent. You have read the information provided above. You have been given the opportunity to ask questions and all the questions have been answered to your satisfaction. You have been provided a copy of this form for your records and you agree to participate in this study. You understand that by agreeing to participate in this research and signing this form, you do not waive any of your legal rights.

☐ I consent to participate in the research study.

☐ I do not consent to participate in the research study.

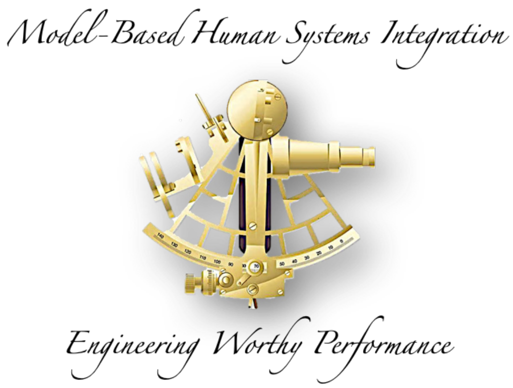
____________________________________ ____________________
Participant's Signature Date

APPENDIX B. MBHSI STUDY VOLUNTEER FLYER



Research Volunteers Needed
Flight Simulation!

Volunteers are needed to participate in an exciting, real-world research project. You will spend about an hour capturing your performance capacity envelope across eight cognitive and psychomotor dimensions in addition to about an hour flying the simulator!



Sign up fast before slots are filled!
1 Oct 2019 to 1 Mar 2020
<u>All volunteers must meet the following criteria:</u>
- At least 18 years old
- Federal Government Employee
- Zero time piloting any aircraft

If interested and qualify, please contact Matt Taranto at mttarant2@nps.edu
This is an NPS IRB APPROVED protocol
For IRB related questions please contact Bryan Hudgens at bjhudgen@nps.edu

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. MBHSI STUDY DEMOGRAPHICS SHEET

Participant Profile Worksheet

Research Title:
System Performance Forecasting for Model-Based Human Systems Integration

Participant Identification Number: _______________ Date: ___________________

Physical Characteristics: (Check all that apply)

Age: _____________

Gender:
- Male
- Female

Color Blindness:
- No
- Protanope (Red Insensitive)
- Deuteranope (Green Insensitive)

Visual Acuity:
- 20/20
- 20/20 Corrected

Hearing:
- Normal
- Impaired
- Corrected

Education: (Highest Level Achieved)

- Didn't graduate high school
- High School Graduate
- Technical Degree
- Undergraduate Degree
- Master's Degree
- Doctoral Degree

Gaming Experience:

- None
- Less than an hour a month
- About an hour a week
- More than an hour a day

Gaming Experience by Type: (Check all that apply)

- Multiplayer Online
- Simulations (Flight)
- Adventure
- Real-Time Strategy
- Puzzle
- Action
- Stealth Shooter
- Combat
- First Person Shooter
- Sports
- Role-Playing
- Educational

Flight Experience by Type: (Check all that apply)

- Ground Training
- Simulator Experience
- Live Flight Training
- Solo Flight
- Licensed Pilot:

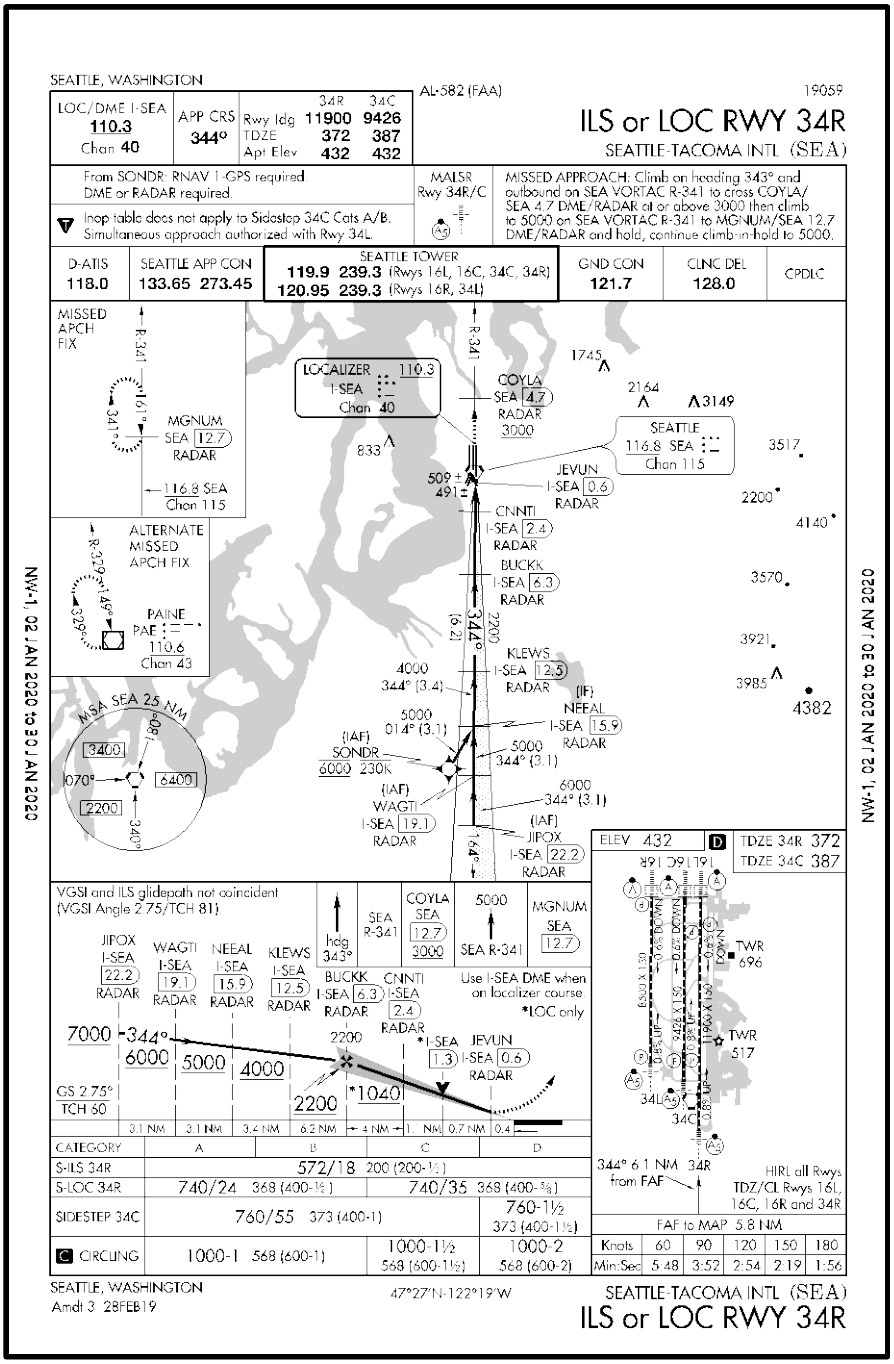
Type: ___________________

Total Hrs: ____________________

Engineering Worthy Performance using General Systems Performance Theory

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D. SEATTLE-TACOMA INTERNATIONAL AIRPORT (KSEA) ILS RWY 34R



THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E. STUDY CHECKLISTS AND PROTOCOL

START-UP

1. Keys-STOWED
2. Lights-ON
3. Fan-OFF
4. Computer-ON
5. TV-ON
6. X-Plane Controls-CENTERED
7. Trash-EMPTY
8. Multi-Choice-ON
9. Bassin Timer-ON
10. Multi-Limb Error Meter-ON
11. Rotary Pursuit OPS Check-COMPLETE
12. ANAM OPS Check-COMPLETE
13. Participant Data Collection Workbook-BUILT
14. Participant ID Number Cross Check-COMPLETE
15. Participant Group Flow-COMPLETE
16. Certificate-PRINTED
17. Folder Review-COMPLETE

PARTICIPANT INSTRUCTIONS

Start-Up Checklist-COMPLETE.

1. Welcome participant-TIME ANNOTATED
2. Restroom-CHECK
3. Cell Phones-OFF
4. Consent Form-SIGNED
5. Demographics-COMPLETE

VERBAL INSTRUCTIONS

INTRODUCTION:
We are about to begin the first phase of the experiment. During this time, you will be completing eight different cognitive and psychomotor tests. For each test, your absolute best effort is requested. Instructions for each test will be given just prior to completing each test.

MULTI-LIMB COORDINATION
For this test your objective is to trace the star with the two black handles in the direction of instruction as quickly as possible. Your score is based on the number of seconds it takes to make a complete star. Each time you exit the black star path, you will be assessed a one-second penalty. So, you are measured on both speed and accuracy. The lower the score the better. Please do not press down on the needle as it will damage the equipment. You will

be given one practice rotation in each direction. Then you will be scored on each direction with the lowest of the two being recorded.

BASSIN-TIMER
Please stand at the end of the runway with the clicker in your dominant hand. An amber light will illuminate for some period of time warning you that the running lights is about to start. You are to press the black button when you think the last light on the runway will illuminate. You will be given two practice runs and five test runs. Your score closest to zero will be recorded. Speeds will vary during the tests.

MULTI-CHOICE REACTION TIMER
Place your dominant index finger on the pink square. There are four lights each with a corresponding button used to de-illuminate each light. The center button is for the buzzer. When you hear a buzzer, press the center button. Return your finger to the square after each test. The order and timing will be random. You will be given two practice runs and five test runs. Your fastest test run will be recorded as your score.

ROTARY-PURSUIT
Place the wand in your dominant hand and over the yellow square. The objective is to keep the wand over the illuminated light that is moving. Please do not touch the glass surface. You will be given two 30-second practice trials, one in each direction. You will then be scored on two 30-second tests, one in each direction. Your score that has the highest time on target ratio for 30 seconds will be recorded.

SPEED OF CLOSURE
You will be given two paper tests, one referred to as snowy pictures and the other as concealed words. You will be allowed to read the instructions and one minute times read to you. Your objective is to get as many correct as possible. Your highest score will be recorded.

ANAM
You will now take the Switching Test in ANAM which measures speed and accuracy for basic math and spatial orientation. All instructions are provided. Scores are provided by ANAM following your test.

INFORMATION TEST
You will complete a simple quiz testing your knowledge of basic aircraft flight information and instrument specific information. If receiving training, you will complete this quiz after training a second time.

The cognitive and psychomotor testing portion of the study is complete.

COMFORT BREAK-5 Min

FLIGHT SIMULATION

TRAINING PARTICIPANTS-WATCH VIDEOS
You will now watch two short videos pertaining to fixed wing flight and basic instrument flight specifics.

Second Paper Quiz-COMPLETE

NON-TRAINING PARTICIPANTS-NO VIDEOS

You will now start the flight simulation portion of the study. Your first flight will start mid-air and provide the opportunity to gain a sense of the controls and their corresponding responses based on inputs. Give special attention to the AIRSPEED indicator, ALTIMETER, ARTIFICIAL HORIZON, and the CDI. I will provide participant coaching during the first approach and answer any questions. During the second approach, I will provide coaching until you reach six miles from the runway, then I will stop. Your objective is to maintain Airspeed of 65kts +- 10kts and vertical/horizontal alignment on the CDI. REMEMBER: CHASE THE NEEDLES. These three performance indicators will be scored every second from six miles to one mile from the runway. Please do your best not to violate the stated and coached performance thresholds of 65kts and 2.5degrees of CDI at all times. For every half degree or knot from center you will be assessed a penalty that is the square of the deviation. So, for three knots, it is a penalty of 9.

QUIZ:

1) What is your target Airspeed?
Answer: 65Kts

2) Where is the CDI and how do you use it to navigate to the runway?
Answer: Right gauge, if needle to the right, flight right, if needle low, descend.

You will fly the approach twice. The first will be your performance score and the other will be to demonstrate learning or performance changes with partial task automation depending on your participant track.

Approach 1-COMPLETE

Approach 1 Data-SAVE AS ID NUMBER

AUTO-PILOT-ON/OFF

Approach 2-COMPLETE

Approach 2 Data-SAVE AS ID NUMBER(2)

Written BPR Data Cross Check-COMPLETE

Simulation Data Cross Check-COMPLETE

Certificate-SIGN

NASA TLX-COMPLETE

Study-COMPLETE

Participant-DONE

Data Analytics-COMPLETE

Data Saved in Sakai-COMPLETE

Data Saved on Computer-COMPLETE

Data Saved in Box-COMPLETE

BPR Collection Tools-OFF

Fan-ON

Monitors-OFF

Lights-OFF

Keys-Check

APPENDIX F. MBHSI STUDY CERTIFICATE OF APPRECIATION

CERTIFICATE OF APPRECIATION

FOR

Participant Name

Model-Based Human Systems Integration

**THANK YOU FOR YOUR RESEARCH PARTICIPATION
IN THE FIRST-EVER MODEL-BASED APPROACH TO
HUMAN SYSTEMS INTEGRATION**

SIMULATED ILS APPROACH TO SEATTLE-TACOMA AIRPORT IN A CESSNA–172 USING X-PLANE 11
&
BASIC PERFORMANCE RESOURCE CAPACITY MEASUREMENT
AT
THE NAVAL POSTGRADUATE SCHOOL

Engineering Worthy Performance

R_A
TWENTY TWO-DIMENSIONAL PERFORMANCE CAPACITY ENVELOPE AS MEASURED WAS:

____

____

Matthew Taranto
Lt Col, USAF, BSC, CAsP
Ph.D. Candidate, MOVES Institute

HLT_P
ILS APPROACH PERFORMANCE AS MEASURED WAS:

____

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX G. STUDY BPR DETAILS

Table 32. Visual Motion Tracking Accuracy

1. BPR: Visual Motion Tracking Accuracy
Fleishman's Ability: Control Precision
Definition: The capacity to make highly controlled and precise adjustments in moving the controls of a vehicle/machine quickly and repeatedly to exact positions.
Priority (Limiting BPR) Rank: 1
Equipment: Rotary Pursuit Test by Lafayette Instrument Co., Inc.
Measurement: Rate = Time on target/30 seconds
Protocol: Place the wand in your dominant hand and over the yellow square. The objective is to keep the wand over the illuminated light that is moving. Please do not touch the glass surface. You will be given two 30-second practice trials, one in each direction. You will then be scored on two 30-second tests, one in each direction. Your score that has the highest time on target ratio for 30 seconds will be recorded.

Table 33. Visual Motion Prediction Accuracy

2. BPR: Visual Motion Prediction Accuracy
Fleishman's Ability: Rate Control
Definition: The capacity to adjust an equipment control in response to changes in the speed and/or direction of a continuously moving object or scene.
Priority (Limiting BPR) Rank: 2
Equipment: Bassin Anticipation Timer by Lafayette Instrument Co, Inc.
Measurement: 1/response error in sec
Protocol: Please stand at the end of the runway with the clicker in your dominant hand. An amber light will illuminate for some period of time warning you that the running lights is about to start. You are to press the black button when you think the last light on the runway will illuminate. You will be given two practice runs and five test runs. Your score closest to zero will be recorded. Speeds will vary during the tests.

Table 34. Multi-limb Coordination

3. BPR: Multi-Limb Coordination Throughput **4. BPR: Multi-Limb Coordination Speed** **5. BPR: Multi-Limb Coordination Accuracy**
Fleishman's Ability: Multi-Limb Coordination
Definition: The capacity to coordinate movements of two or more limbs while moving controls.
Priority (Limiting BPR) Rank: 3
Equipment: Two Arm Coordination Test by Lafayette Instrument Co, Inc.
Measurements: o Throughput = (Speed*Accuracy) o Speed was measured in cm/sec (track was 48cm) o Accuracy was measured by using an exponential penalty = EXP(-0.065*number of penalties)*speed
Protocol: For this test, your objective is to trace the star with the two black handles in the direction of instruction as quickly as possible. Your score is the speed at which you complete a star. Each time you exit the black star path, you will be assessed a penalty. So, you are measured on both speed and accuracy. The higher the score, the better. Please do not press down on the needle as it will damage the equipment. You will be given one practice rotation in each direction. Then you will be scored on each direction with the fastest of the two being recorded.

Table 35. Multi-choice Reaction Speed

6. BPR: Multi-Choice Reaction Speed
Fleishman's Ability: Response Orientation
Definition: The capacity to choose between two or more movements quickly and correctly when two or more different signals are given.
Priority (Limiting BPR) Rank: 4
Equipment: Multi-Choice Reaction Timer by Lafayette Instrument Co, Inc.
Measurement: 1/response error in sec
Protocol: Place your dominant index finger on the pink square. There are four lights, each with a corresponding button used to de-illuminate each light. The center button is for the buzzer. When you hear a buzzer, press the center button. Return your finger to the square after each test. The order and timing will be random. You will be given two practice runs and five test runs. Your fastest test run will be recorded as your score.

Table 36. Spatial Orientation

7. BPR: Spatial Orientation Response Throughput **8. BPR: Spatial Orientation Response Speed** **9. BPR: Spatial Orientation Response Accuracy**
Fleishman's Ability: Spatial Orientation
Definition: The capacity to know one's location in relation to the environment
Priority (Limiting BPR) Rank: 5
Equipment: Automated Neuropsychological Assessment Metrics (ANAM) Switching Task (spatial orientation)
Measurements: o Throughput = 60,000/response time in milliseconds*Accuracy o Speed = 60,000/response time in milliseconds o Accuracy = number correct/total
Protocol: Provided by ANAM.

Table 37. Perceptual Integration Capacity

10. BPR: Perceptual Integration Capacity (Concealed Words) **11. BPR: Perceptual Integration Capacity (Snowy Pictures)**
Fleishman's Ability: Speed of Closure
Definition: The capacity to quickly make sense of information that initially seems to be without meaning or organization.
Priority (Limiting BPR) Rank: 6
Equipment: Educational Testing Services—Concealed Words Test (CS-2) Educational Testing Services—Snowy Pictures Test (CS-3)
Measurement: Percent Correct
Protocol: Provided by Educational Testing Services. Time for CS-2 was adjusted to four minutes for all 50 words Time for CS-3 was adjusted to three minutes for all 24 snowy pictures

Table 38. Math Processing

12. BPR: Math Processing Throughput **13. BPR: Math Processing Speed** **14. BPR: Math Processing Accuracy**
Fleishman's Ability: N/A
Definition: The capacity to quickly and accurately process numerical information
Priority (Limiting BPR) Rank: N/A
Equipment: ANAM Switching Task (math)
Measurements: o Throughput = 60,000/response time in milliseconds*Accuracy o Speed = 60,000/response time in milliseconds o Accuracy = number correct/total
Protocol: Provided by ANAM.

Table 39. Switching Task

15. BPR: Switching Response Throughput **16. BPR: Switching Response Speed** **17. BPR: Switching Response Accuracy**	
	Fleishman's Ability: N/A
	Definition: The capacity to quickly and accurately switch between tasks
	Priority (Limiting BPR) Rank: N/A
	Equipment: ANAM Switching Task
Measurements:	○ Throughput = 60,000/response time in milliseconds*Accuracy ○ Speed = 60,000/response time in milliseconds ○ Accuracy = number correct/total
	Protocol: Provided by ANAM.

Table 40. Knowledge BPR

18. BPR: Basic Flight Knowledge **19. BPR: Basic Aircraft Instrument Knowledge**	
	Fleishman's Ability: N/A
	Definition: The capacity of basic flight and aircraft instrument knowledge
	Priority (Limiting BPR) Rank: N/A
	Equipment: Multiple choice, fill in the blank, and True/False tests developed by the researcher derived from FAA Exam Test Questions.
Measurement: Percent Correct	
	Protocol: No time limit. All tests were taken using paper and pencil.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX H. BASIC AIRCRAFT INFORMATION AND INSTRUMENTATION KNOWLEDGE TESTS

Model-Based Human Systems Integration

Engineering Worthy Performance

ILS Information BPR Test

August 2019

Flight Characteristics

1. The tail section of an airplane is known as the:

A. Aft Section
B. Stern
C. Empanage
D. Tail

2. The Ailerons on an aircraft are designed to change:

A. Pitch or Altitude
B. Roll or Direction
C. Yaw or Heading
D. Steering direction during ground operations

3. The CDI is known as:

A. Course Direction Indicator
B. Certified Deviance Instrument
C. Complex Direction Indicator
D. Course Deviation Indicator

4. The elevators on an aircraft change:

A. Roll or Direction
B. Yaw or Heading
C. Pitch or Altitude
D. Steering direction during ground operations

5. What does ILS stand for?

____

6. What is Decision Height in terms of an ILS Approach?

____

7. An ILS glide slope is a 5 degree slope to the runway from altitude, True or False?

____

Instrument Approaches

1. What are the typical landing minimums for a CAT I ILS Approach with all components operative?

A. Visibility—2,400 RVR or ½ statute mile; DH—200ft Mean Sea Level
B. Visibility—1,200 RVR or ½ statute mile; DH—200ft above touch down zone elevation
C. Visibility—2,400 RVR or ½ statute mile; DH—200ft above touch down zone elevation

2. What is the full-scale deflection of a CDI when tuned to a localizer?

A. 10 degrees
B. 5 degrees
C. 2.5 degrees

3. Approximately what height is the glide slope centerline at the Middle Marker (MM) of a typical ILS approach?
A. 100 ft
B. 200 ft
C. 300 ft

4. If during an ILS approach in Instrument Flight Rules (IFR) conditions, the approach lights are not visible upon arrival to the Decision Height (DH), the pilot is…

A. Required to immediately execute the missed approach procedure.
B. Permitted to continue approach and descend to the localizer Minimum Descent Altitude (MDA).
C. Permitted to continue approach to the approach threshold of the ILS runway.

5. Which indication will a pilot receive where an Outer Marker (OM) is installed on a front Course ILS approach?

A. Amber light
B. Blue light
C. White light
D. Green light

1

2

3

4

6. Which CDI indications confirm you are right of Course and below glide slope?

A. 1
B. 2
C. 3
D. 4

APPENDIX I. AIRSPEED ANALYTIC APPROACH

- Target: 65kts
- Performance window: 52.5kts—77.5kts (25% more than FAA Standards)
- Actual Airspeed captured once per second
    - All performance data captured to the tenth of a knot
    - Ex: 65.2kts
- Error in kts calculated
    - Observed performance data each second was subtracted from the target of 65kts, absolute value (ABS) ensured positivity.
    - Ex: =ABS(C2-65)
- Performance bins:
    - 1.25kts each for 10 bins either side of target Airspeed
    - If logic used in Excel: =IF(D2>=11.25,"10",IF(D2>=10,"9",IF(D2>=8.75,"8",IF(D2>=7.5,"7",IF(D2>=6.25,"6",IF(D2>=5,"5",IF(D2>=3.75,"4",IF(D2>=2.5,"3",IF(D2>=1.25,"2",IF(D2>=0,"1"))))))))))
    - Ex: 1.25kts-2.5kts = Error of 2
- Error squared
    - Error of 2 = Penalty of 4
- Score: average of penalties spanning HLT
    - Penalties assigned in a separate column were summed then averaged
    - Ex: =AVERAGE(F2:F278)
- Penalties: assessed if Airspeed exceeded the performance window established. If not observed, a score of 1 ensured no influence on final score.
    - Airspeed w/Penalty = EXP(0.05*Number of Penalties)*original Airspeed Score
    - Pass/Fail criteria: =IF(Sum of penalties <=10,"PASS," "FAIL")
- Lower scores equated to better results, but for the purposes of GSPT, better scores need to be higher. So, inverting the score and multiplying it by 100 achieved this outcome.
    - Ex: (1/Airspeed)*100.

Figure 108. X-Plane 11 Cessna-172 Airspeed Indicator

Table 41. Sample Layout of Airspeed HLTp Data Measurement

Time	Target Indicator Deviation	Actual	ABS Deviation	Deviation Multiplier	Penalty Assessed	Violation
1	52.5 to 77.5 Kts	68.7	3.664	3	9.00	0
2	52.5 to 77.5 Kts	68.4	3.405	3	9.00	0
3	52.5 to 77.5 Kts	67.0	1.984	2	4.00	0
4	52.5 to 77.5 Kts	65.3	0.251	1	1.00	0
5	52.5 to 77.5 Kts	63.7	1.298	2	4.00	0
6	52.5 to 77.5 Kts	62.3	2.723	3	9.00	0
7	52.5 to 77.5 Kts	60.8	4.178	4	16.00	0
8	52.5 to 77.5 Kts	59.7	5.338	5	25.00	0
9	52.5 to 77.5 Kts	58.8	6.208	5	25.00	0
10	52.5 to 77.5 Kts	58.1	6.889	6	36.00	0

Table 42. Airspeed Performance Data Example

273	52.5 to 77.5 Kts	65.7	0.726	1	1.00	0
274	52.5 to 77.5 Kts	66.2	1.230	1	1.00	0
275	52.5 to 77.5 Kts	65.7	0.718	1	1.00	0
276	52.5 to 77.5 Kts	64.1	0.906	1	1.00	0
277	52.5 to 77.5 Kts	62.2	2.820	3	9.00	0
AVG			4.62		25.57	0
					Violation Penalty	
					1	
					Adjusted Score	
					25.57400722	
					Pass/Fail	
					PASS	

APPENDIX J. COURSE (CDI-H) AND GLIDESLOPE (CDI-V) ANALYTIC APPROACH

- Target: 0 degrees (deg) of deflection
- Performance window: < full deflection (< 2.5 deg or 0—2.499 deg), (25% more than FAA Standards)
- Actual deflection each second listed in decimal degrees
    - All performance data captured to the thousandth of a degree
    - Ex: .075 deg
- Error in degrees calculated
    - CDI-H or V deflection was recorded each second. ABS ensured positive values.
    - Ex: =ABS(C2)
    - Ex: = -1.249 was converted to 1.249
- Performance bins:
    - .25 deg each for 10 bins either side of target CDI-H or V
    - If logic used in Excel: =IF(D2>=2.25,"10",IF(D2>=2,"9",IF(D2>=1.75,"8",IF(D2>=1.5,"7",IF(D2>=1.25,"6",IF(D2>=1,"5",IF(D2>=0.75,"4",IF(D2>=0.5,"3",IF(D2>=0.25,"2",IF(D2>=0,"1"))))))))))
    - Ex: .25 deg—.5 deg resulted in an error of 2
- Error squared
    - Error of 2 resulted in a penalty of 4
- Score is average of penalties spanning approach
    - Penalties assigned in a separate column, summed, and averaged
    - Ex: =AVERAGE(F2:F278)
- Penalties: assessed if deflection exceeded established performance window. If not observed, a score of 1 ensured no influence on final score.
    - CDI-H or V w/Penalty = EXP(0.05*Number of Penalties)*original CDI-H or V Score
    - Pass/Fail criteria: =IF(Sum of penalties <=10,"PASS," "FAIL")
- Lower scores equated to better results, but for the purposes of GSPT, better scores need to be higher. So, inverting the score and multiplying it by 100 achieved this outcome.
    - Ex: (1/CDI-H or V)*100

Figure 48 provides an example of the CDI. The vertical needle informs Course deviation and the horizontal needle informs Glideslope deviation. Each of the white dots indicates .5

deg deflection. In this example the vertical needle indicates the aircraft is left of Course by at least 2.5 deg and is on Glideslope.

Figure 109. X-Plane 11 Cessna-172 Course CDI

Table 43. Sample Layout of CDI-H or CDI-V HLTp Data Measurement

Time	Target Indicator Deviation	Actual	ABS Actual	Deviation Multiplier	Penalty Assessed	Violation
1	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
2	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
3	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
4	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
5	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
6	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
7	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
8	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
9	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
10	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0

Table 44. CDI-H or CDI-V Performance Data Example

290	**-2.5 to 2.5 Deg**	**1.000**	**1.000**	**5**	**25.00**	**0**
291	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
292	**-2.5 to 2.5 Deg**	**1.000**	**1.000**	**5**	**25.00**	**0**
293	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
294	**-2.5 to 2.5 Deg**	**1.000**	**1.000**	**5**	**25.00**	**0**
295	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
296	**-2.5 to 2.5 Deg**	**1.000**	**1.000**	**5**	**25.00**	**0**
297	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
298	**-2.5 to 2.5 Deg**	**1.000**	**1.000**	**5**	**25.00**	**0**
299	-2.5 to 2.5 Deg	1.000	1.000	5	25.00	0
300	**-2.5 to 2.5 Deg**	**1.000**	**1.000**	**5**	**25.00**	**0**
AVG					25.00	0
					Violation Penalty	
					1	
					Adjusted Score	
					25	
					Pass/Fail	
					PASS	

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX K. MBHSI MODEL IDENTIFIERS

The R-generated models include important identification information at the top of each model. The following reference lists help the reader interpret each model.

Each BPR has a numerical identifier. This list identifies BPRs used in the analysis:

- BPR 1 = Multi-Limb Coordination Speed
- BPR 2 = Multi-Limb Coordination Accuracy
- BPR 3 = Multi-Limb Coordination Throughput
- BPR 4 = Visual Motor Prediction Accuracy
- BPR 5 = Visual Motor Tracking Accuracy
- BPR 6 = Multi-Choice Reaction Speed
- BPR 7 = ANAM Math Speed
- BPR 8 = ANAM Math Accuracy
- BPR 9 = ANAM Math Throughput
- BPR 10 = ANAM Spatial Orientation Speed
- BPR 11 = ANAM Spatial Orientation Accuracy
- BPR 12 = ANAM Spatial Orientation Throughput
- BPR 13 = ANAM Switching Speed
- BPR 14 = ANAM Switching Accuracy
- BPR 15 = ANAM Switching Throughput
- BPR 16 = Perceptual Integration Capacity (Concealed Words)
- BPR 17 = Perceptual Integration Capacity (Snowy Pictures)
- Each HLT/sub-HLT included in Project I uses a vector and numerical identifier:
- V41 = HLTp 1 w/o penalties
- V43 = HLTp 1 w/penalties
- V44 = Airspeed sub-HLTp 1 w/o penalties
- V47 = Airspeed sub-HLTp 1 w/penalties
- V49 = Course (CDI-H) sub-HLTp 1 w/o penalties
- V52 = Course (CDI-H) sub-HLTp 1 w/penalties
- V54 = Glideslope (CDI-V) sub-HLTp 1 w/o penalties
- V57 = Glideslope (CDI-V) sub-HLTp 1 w/penalties

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX L. COMPLETE MBHSI PROJECT I MODEL SET

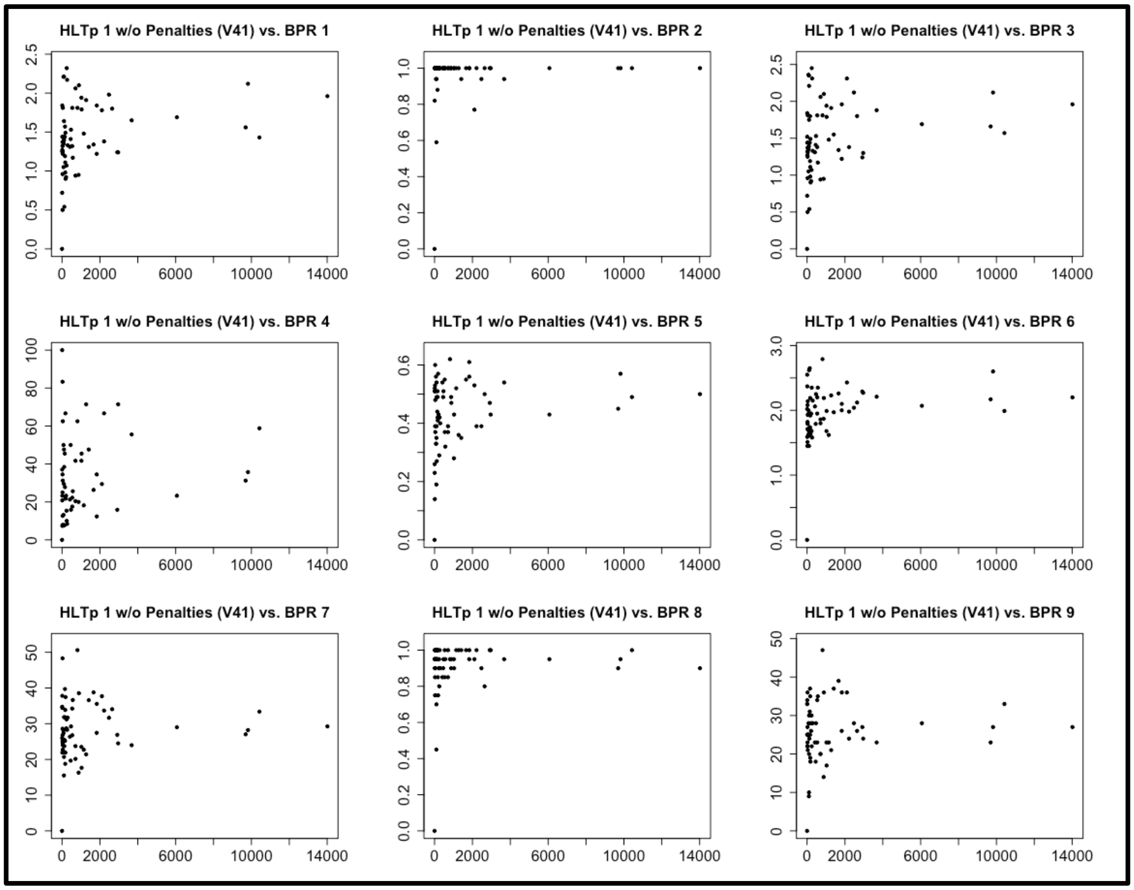


Figure 110. HLTp 1 without Penalties Enforced for BPRs 1–9

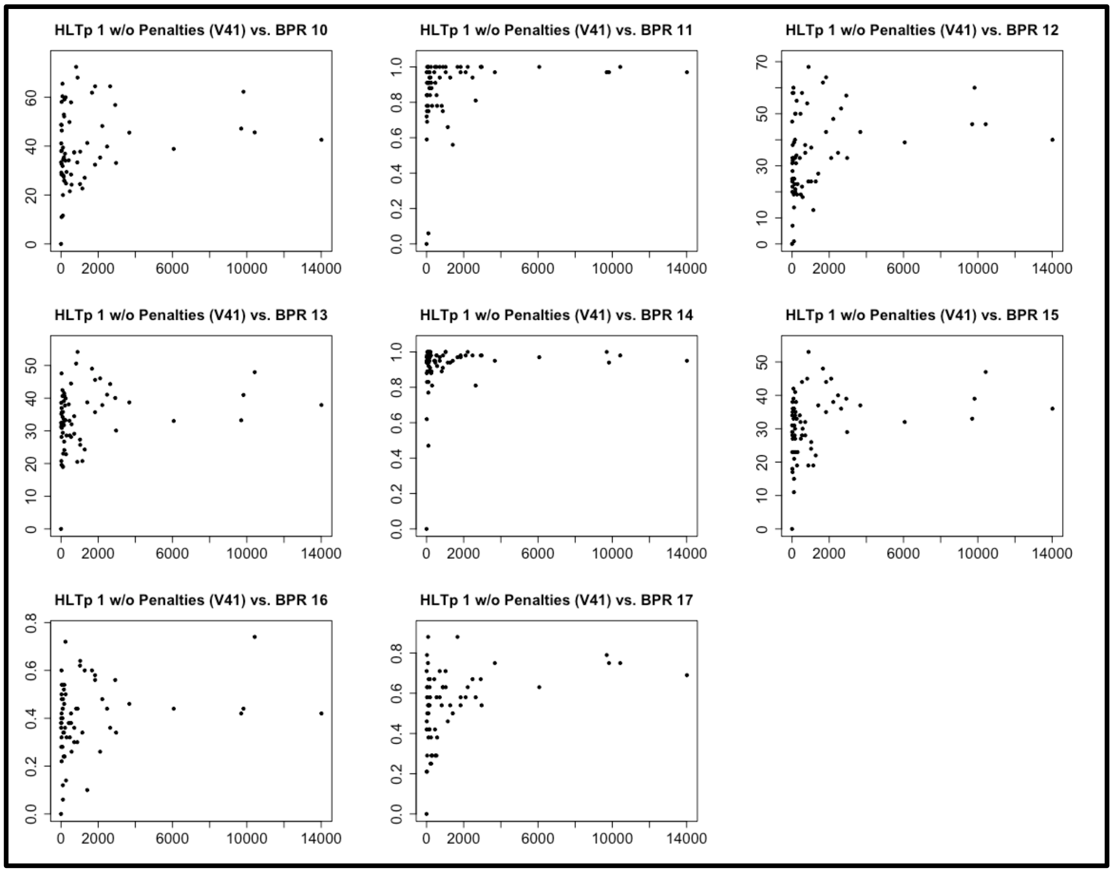


Figure 111. HLTp 1 without Penalties Enforced for BPRs 10–17

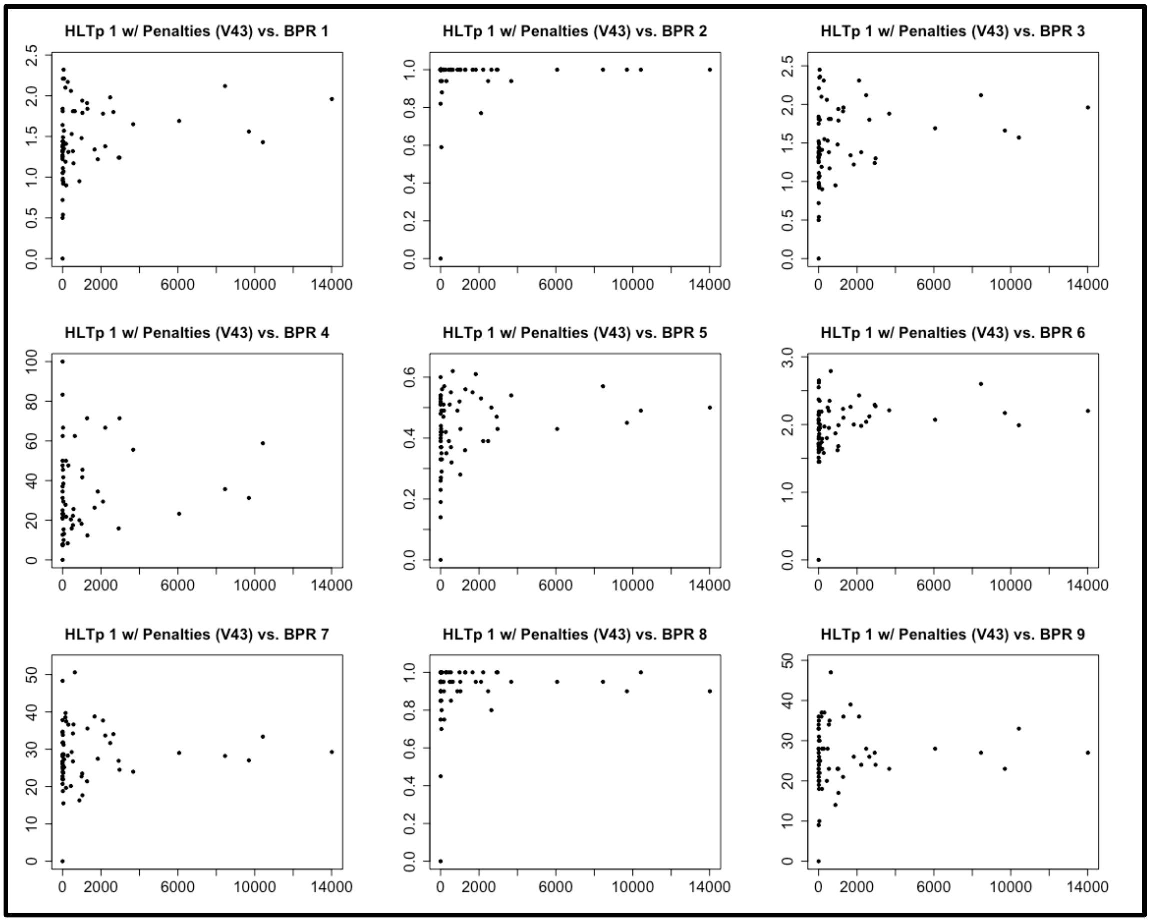


Figure 112. HLTp 1 with Penalties Enforced for BPRs 1–9

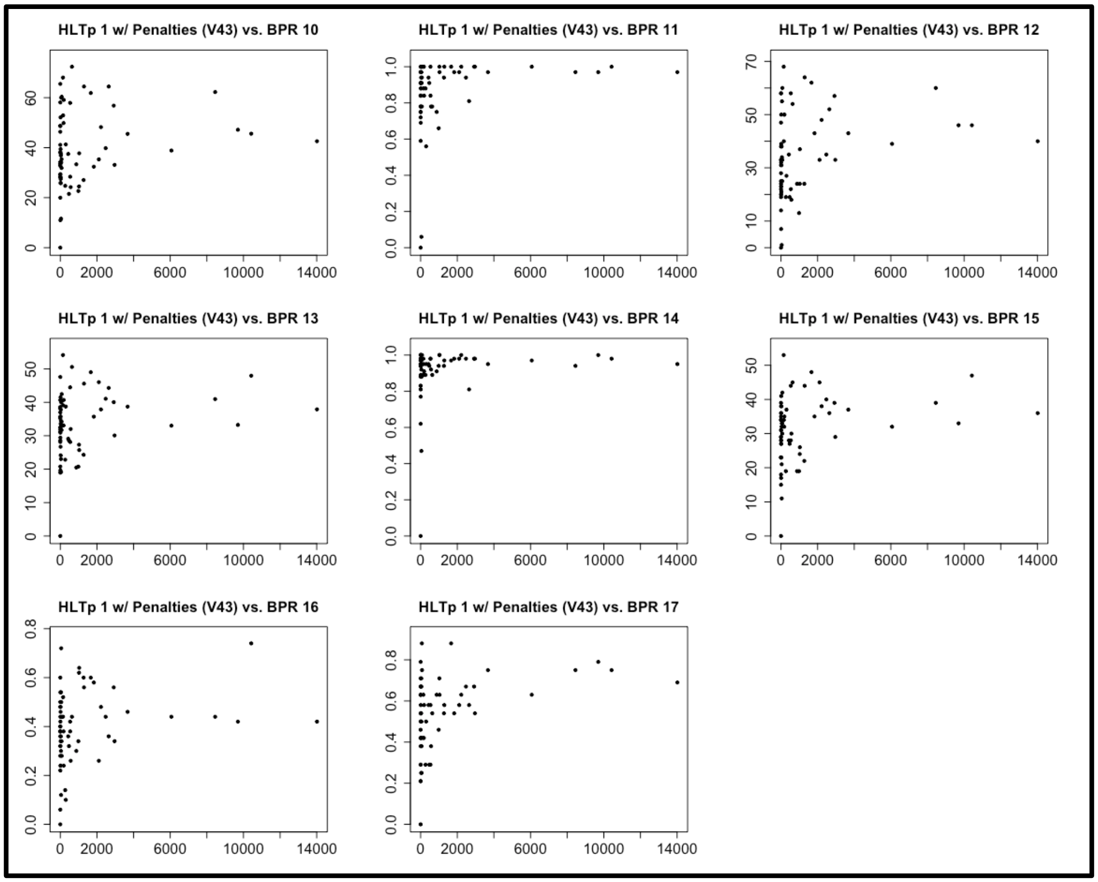


Figure 113. HLTp 1 with Penalties Enforced for BPRs 10–17

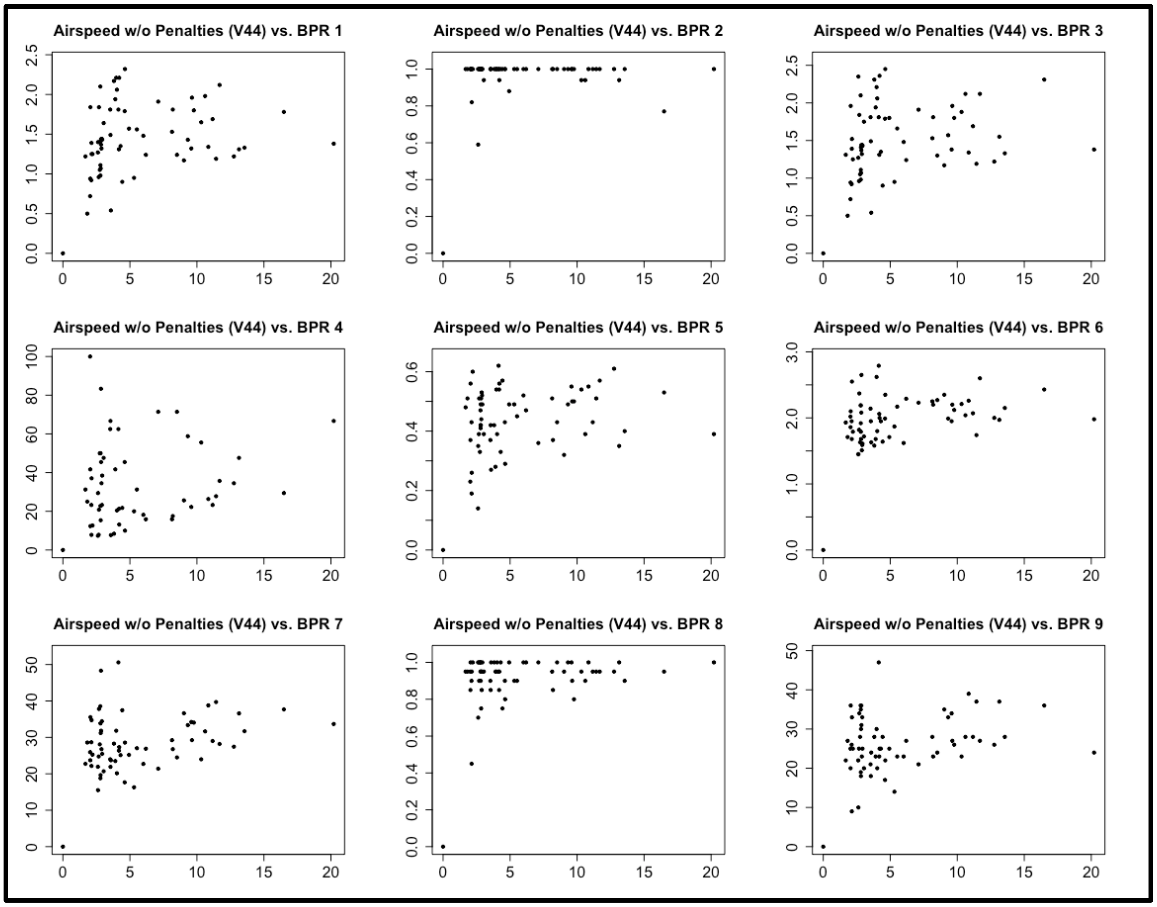


Figure 114. Airspeed models (Sub-HLTp 1) without Penalties Enforced for BPRs 1–9

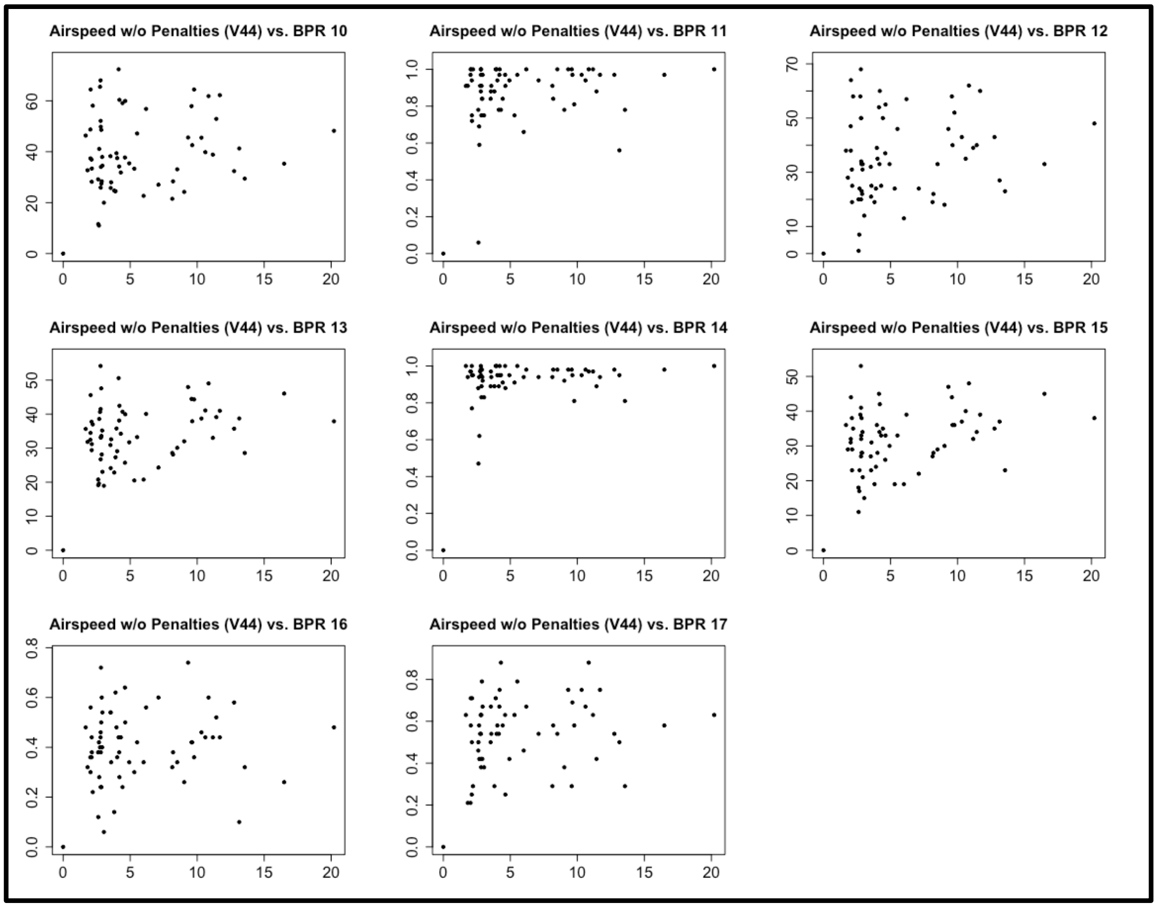


Figure 115. Airspeed (Sub-HLTp 1) without Penalties Enforced for BPRs 10–17

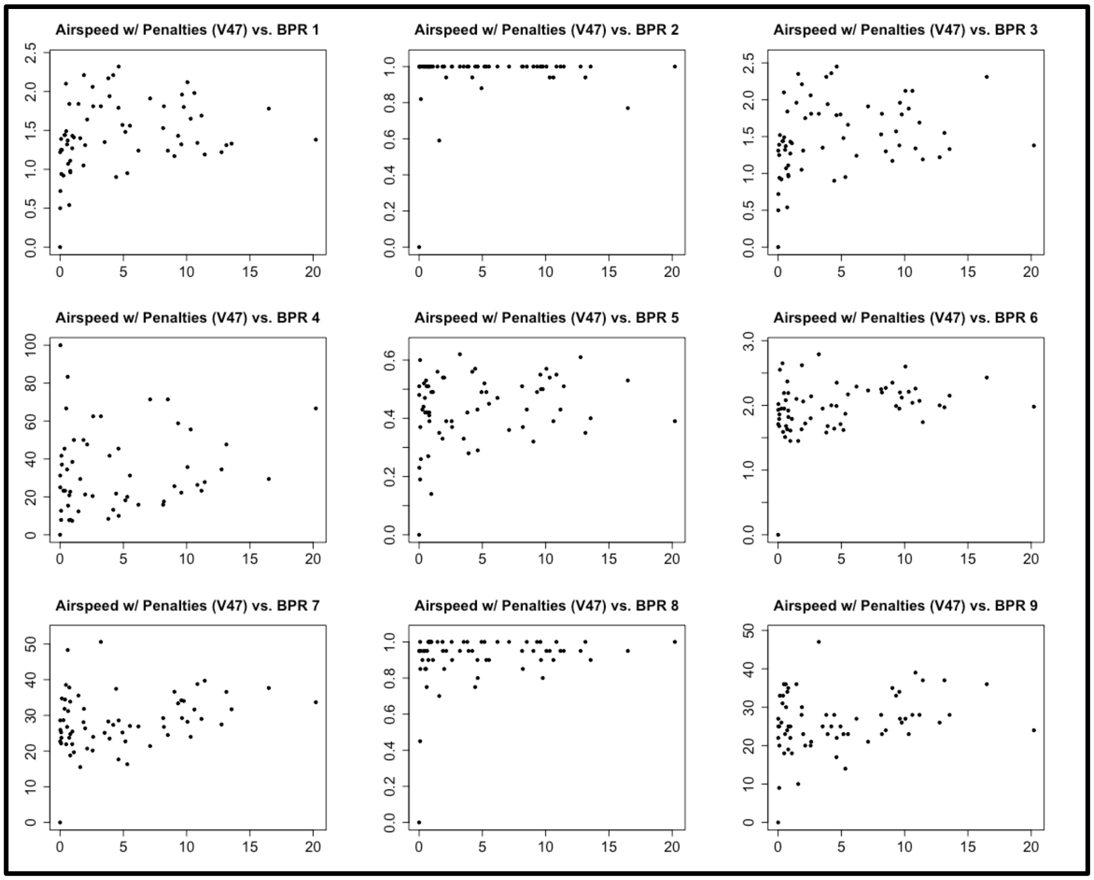


Figure 116. Airspeed (Sub-HLTp 1) with Penalties Enforced for BPRs 1–9

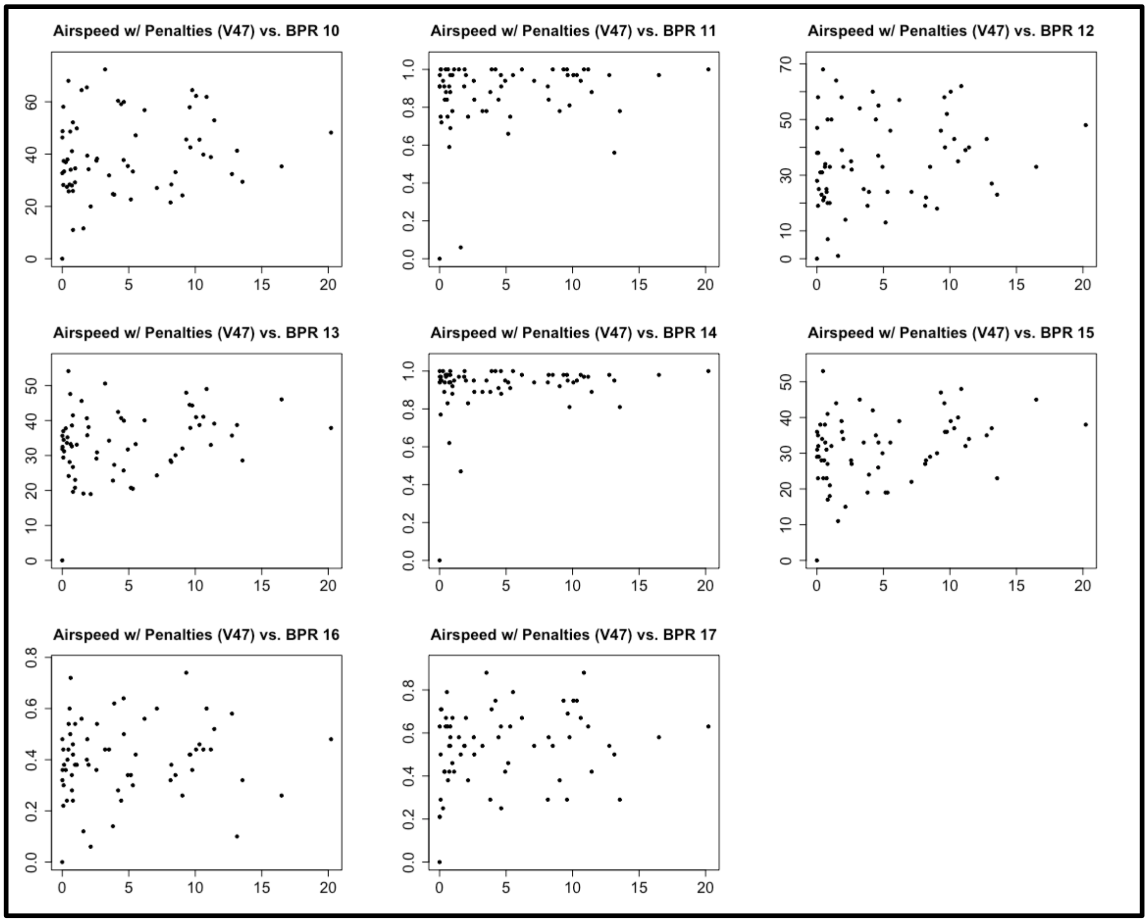


Figure 117. Airspeed (Sub-HLTp 1) with Penalties Enforced for BPRs 10–17

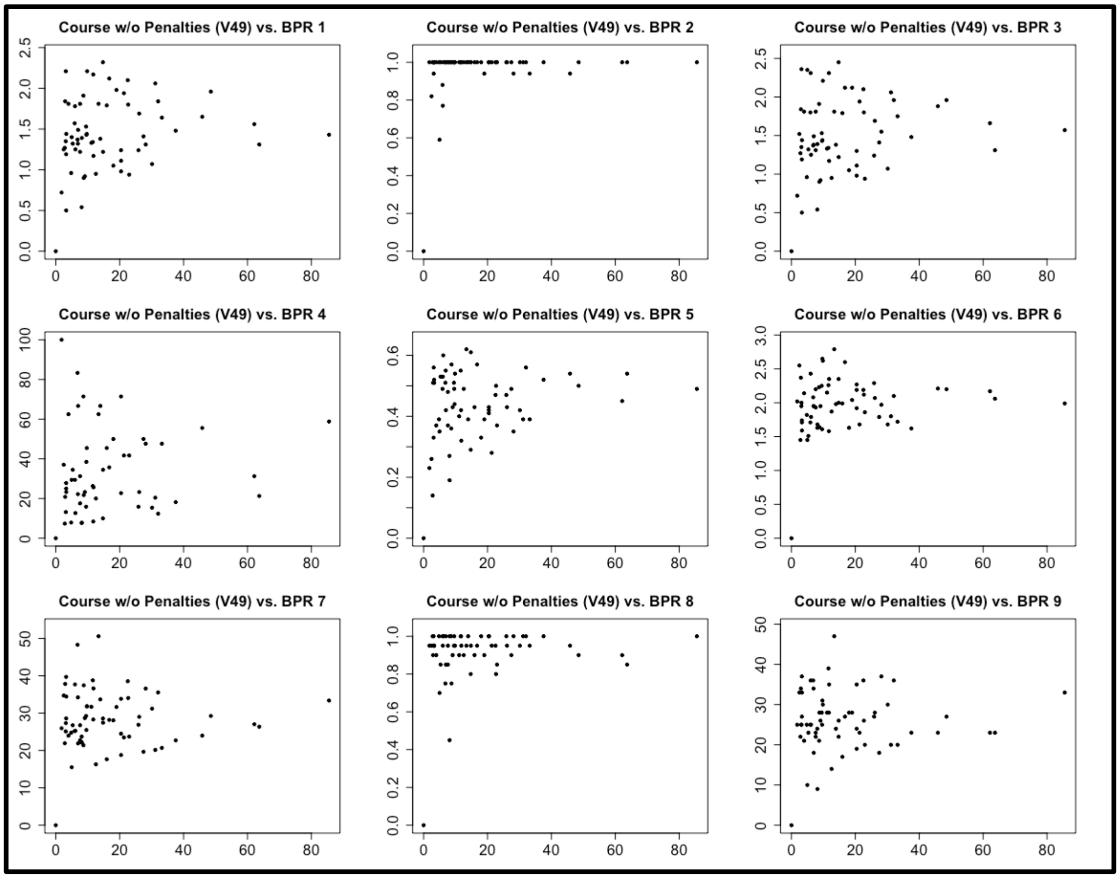


Figure 118. Course (Sub-HLTp 1) without Penalties Enforced for BPRs 1–9

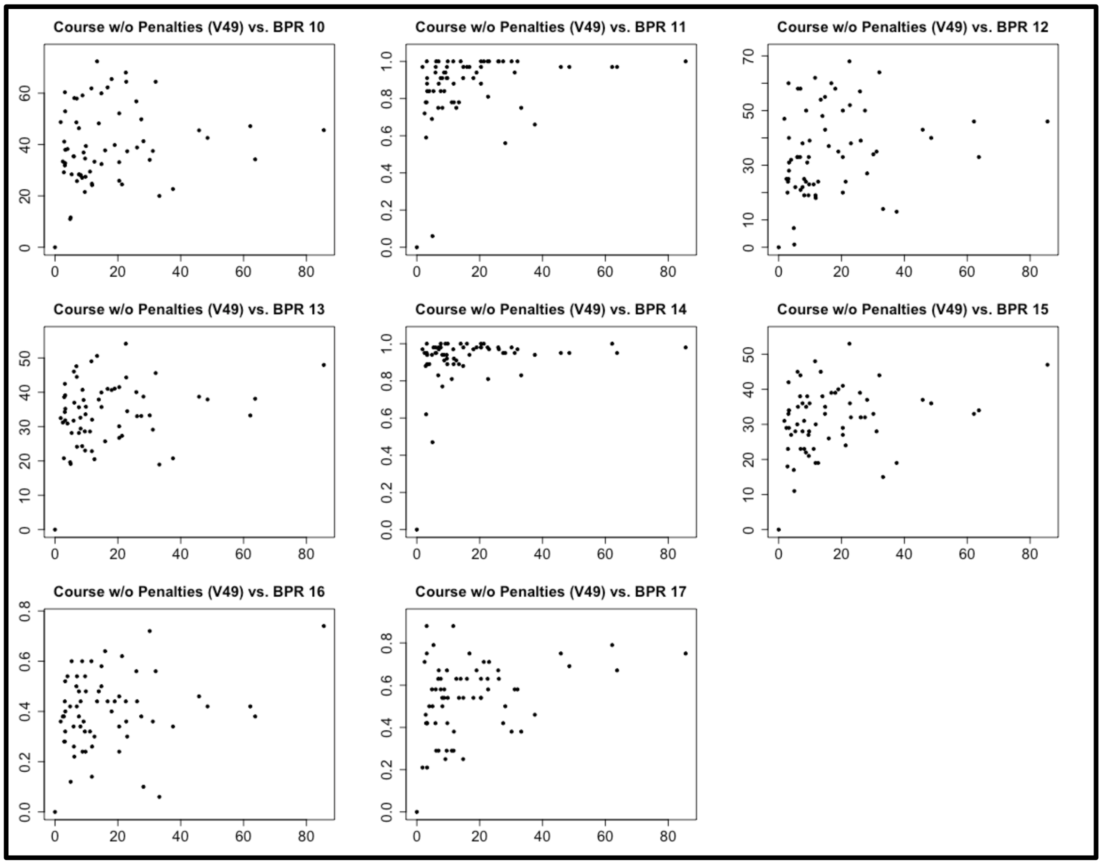


Figure 119. Course (Sub-HLTp 1) without Penalties Enforced for BPRs 10–17

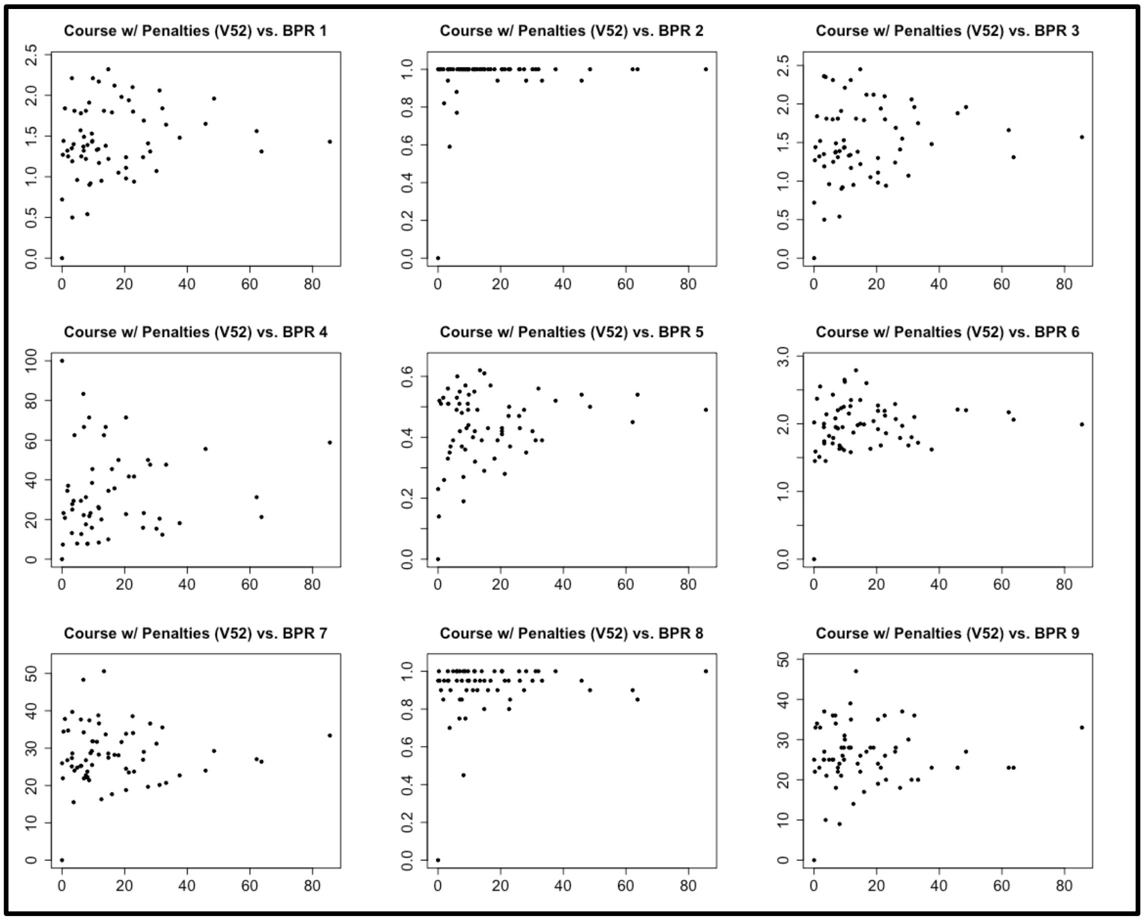


Figure 120. Course (Sub-HLTp 1) with Penalties Enforced for BPRs 1–9

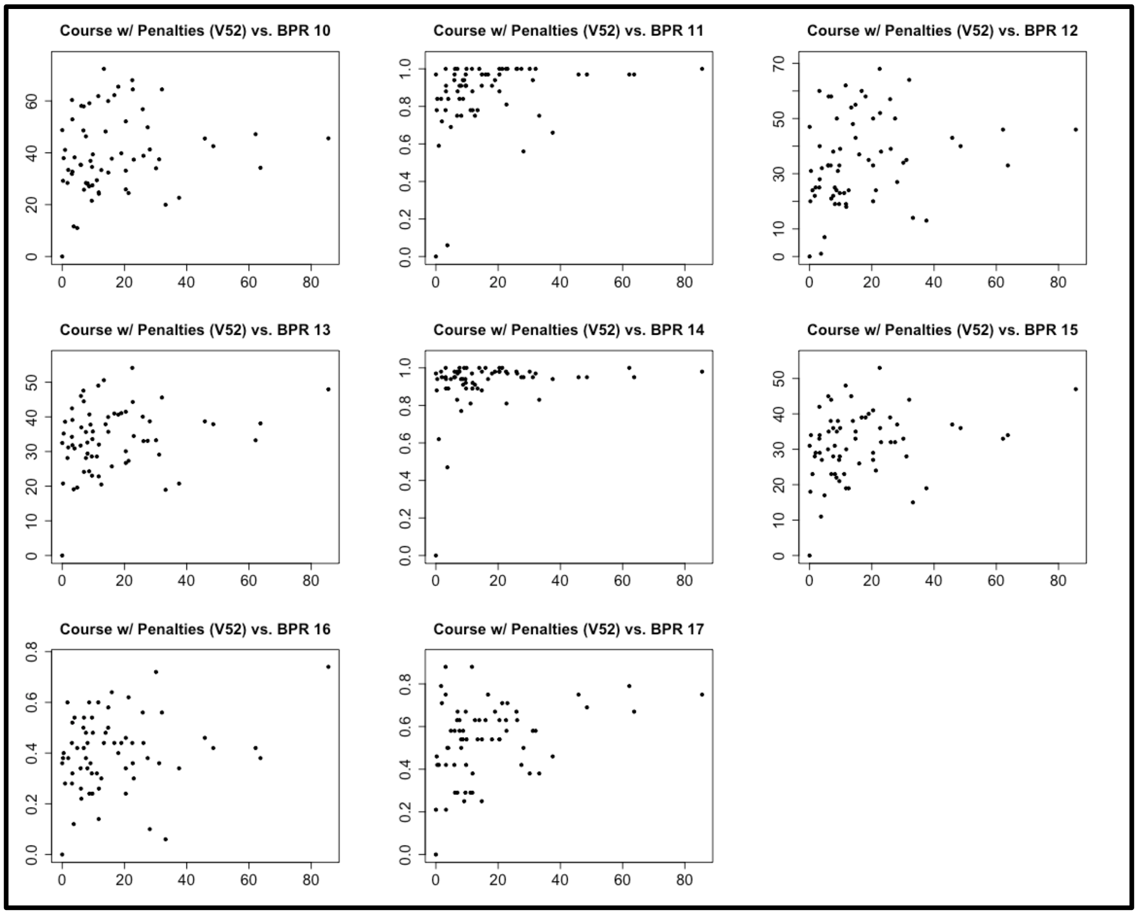


Figure 121. Course (Sub-HLTp 1) with Penalties Enforced for BPRs 10–17

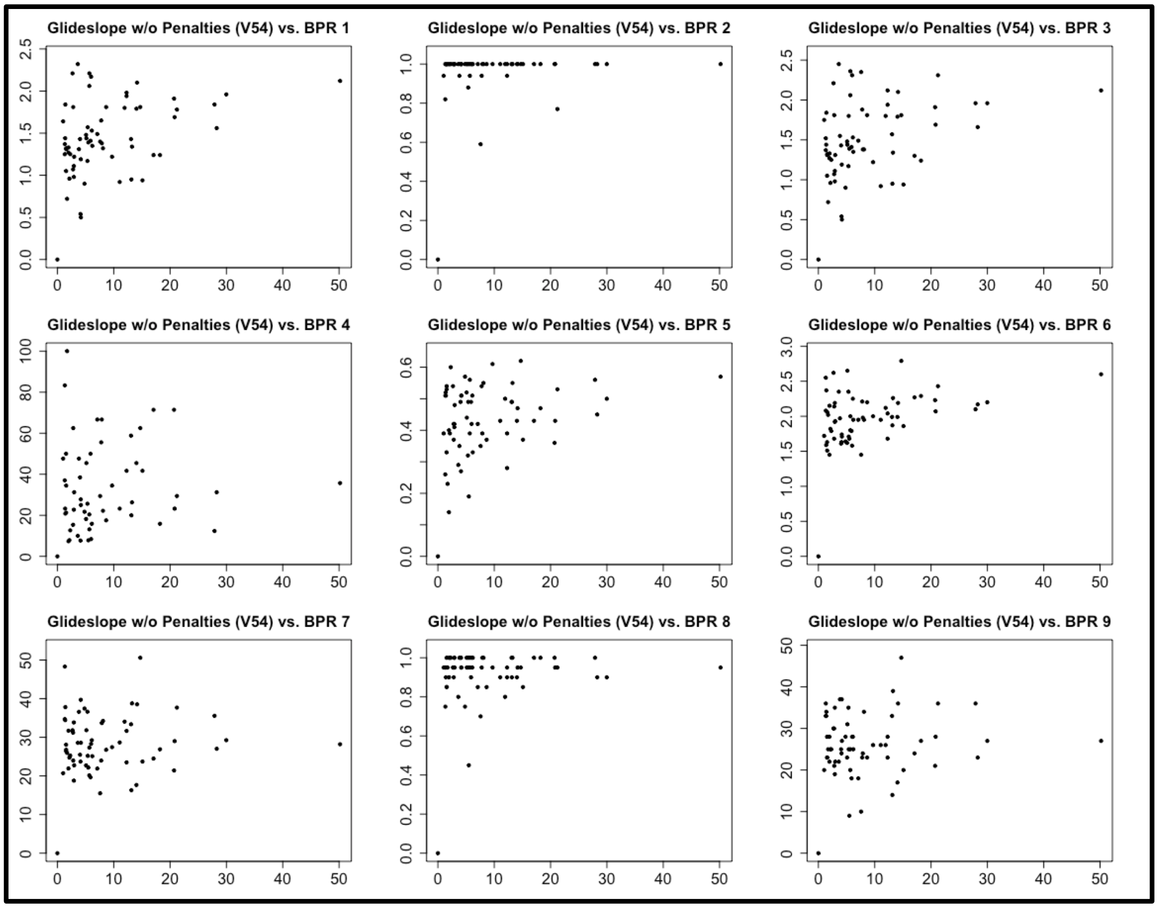


Figure 122. Glideslope (Sub-HLTp 1) without Penalties Enforced for BPRs 1–9

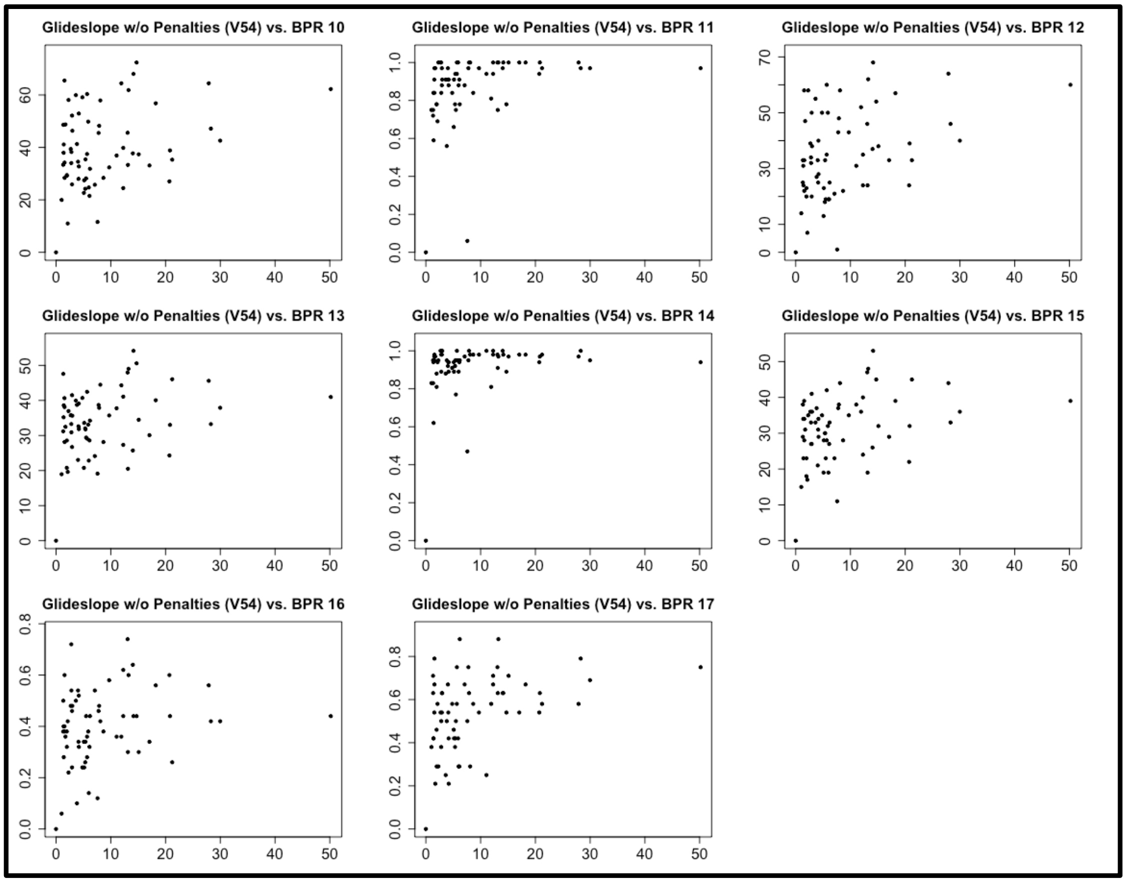


Figure 123. Glideslope (Sub-HLTp 1) without Penalties Enforced for BPRs 10–17

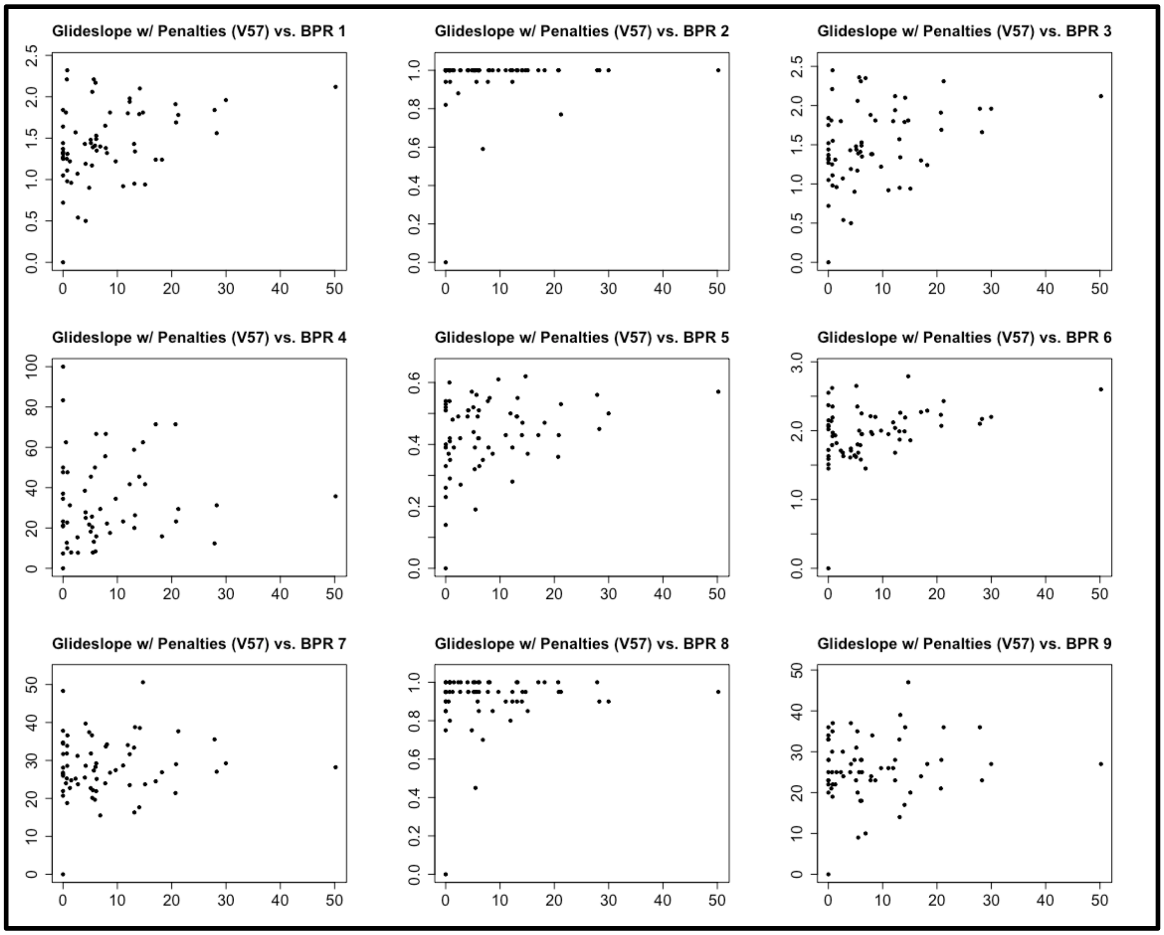


Figure 124. Glideslope (Sub-HLTp 1) with Penalties Enforced for BPRs 1–9

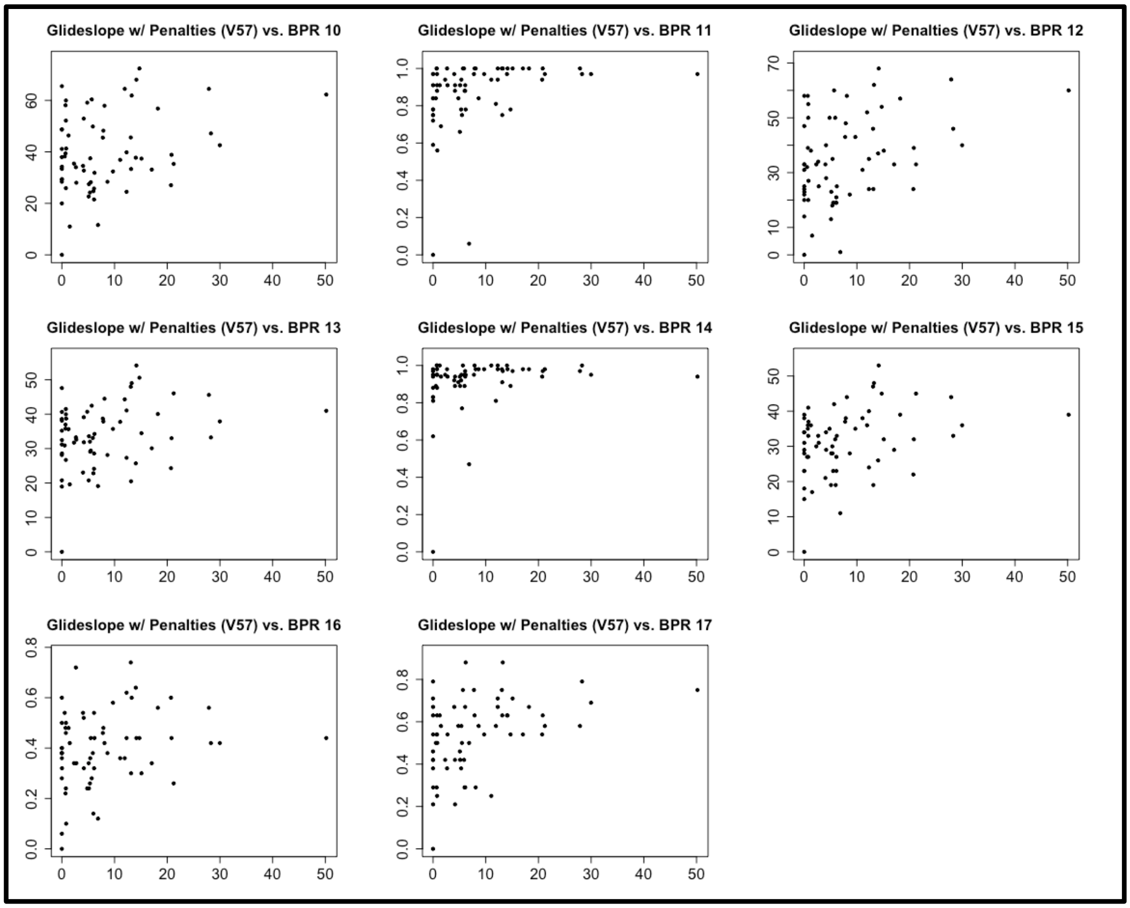


Figure 125. Glideslope (Sub-HLTp 1) with Penalties Enforced for BPRs 10–17

APPENDIX M. PROJECT III BOXPLOTS

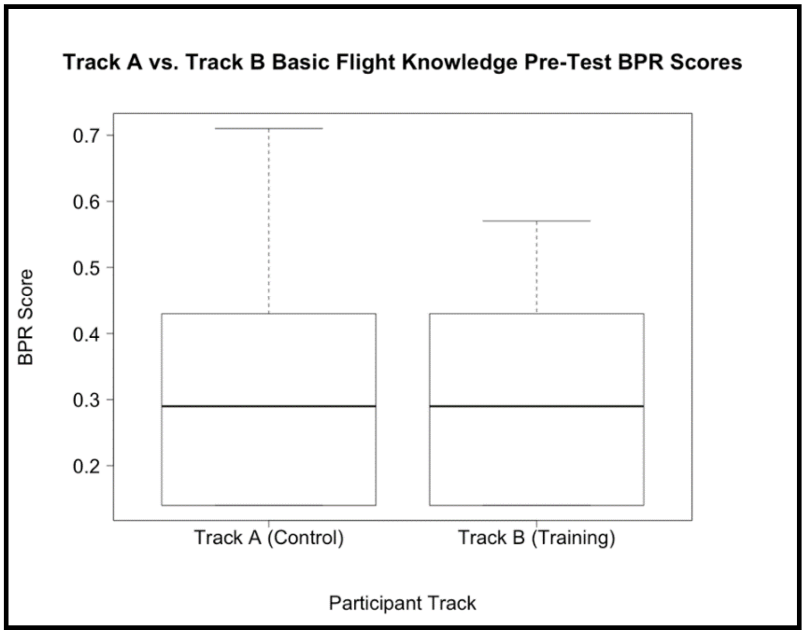


Figure 126. Track A vs. Track B Basic Flight Knowledge Pre-test BPR Scores

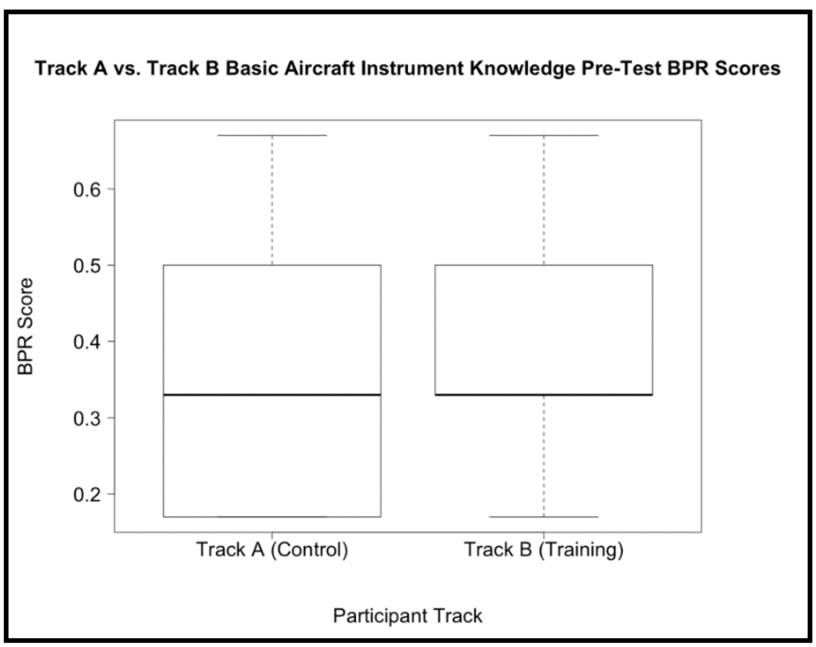


Figure 127. Track A vs. Track B Basic Aircraft Instrument Knowledge Pre-test BPR Scores

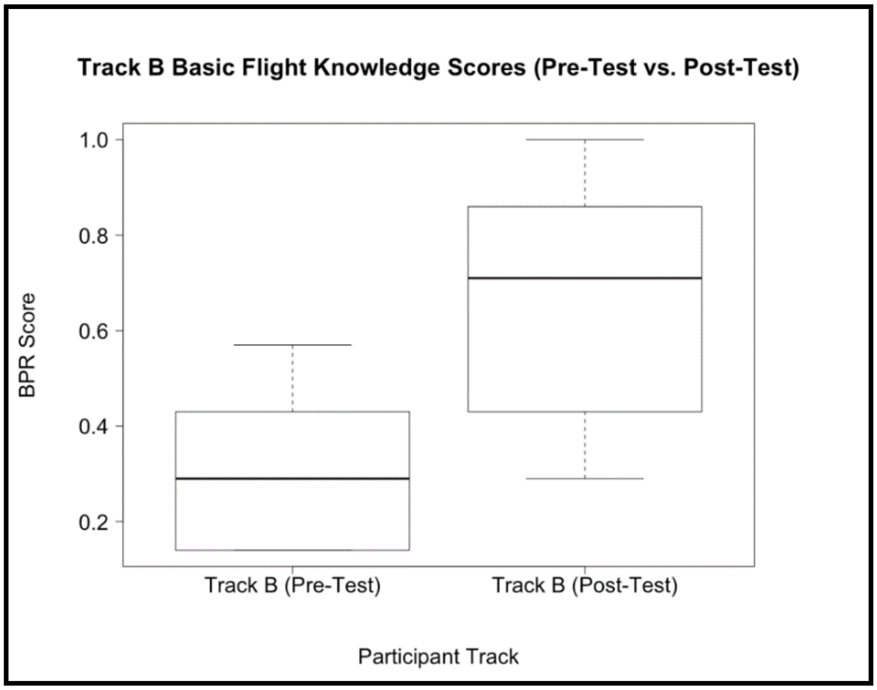


Figure 128. Track B Basic Flight Knowledge Scores (Pre-test vs. Post-test).

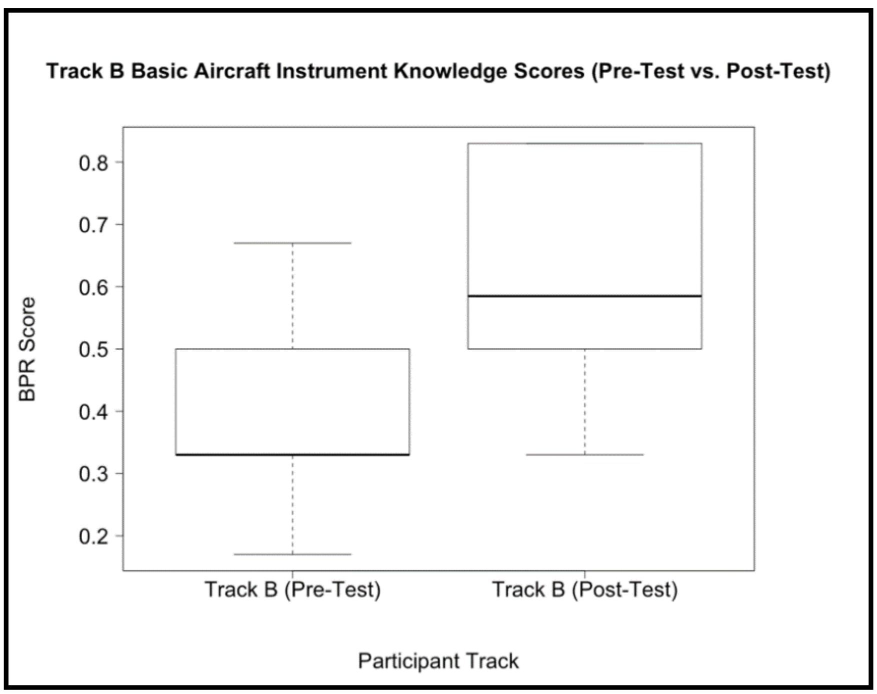


Figure 129. Track B Basic Aircraft Instrument Knowledge Scores (Pre-test vs. Post-test).

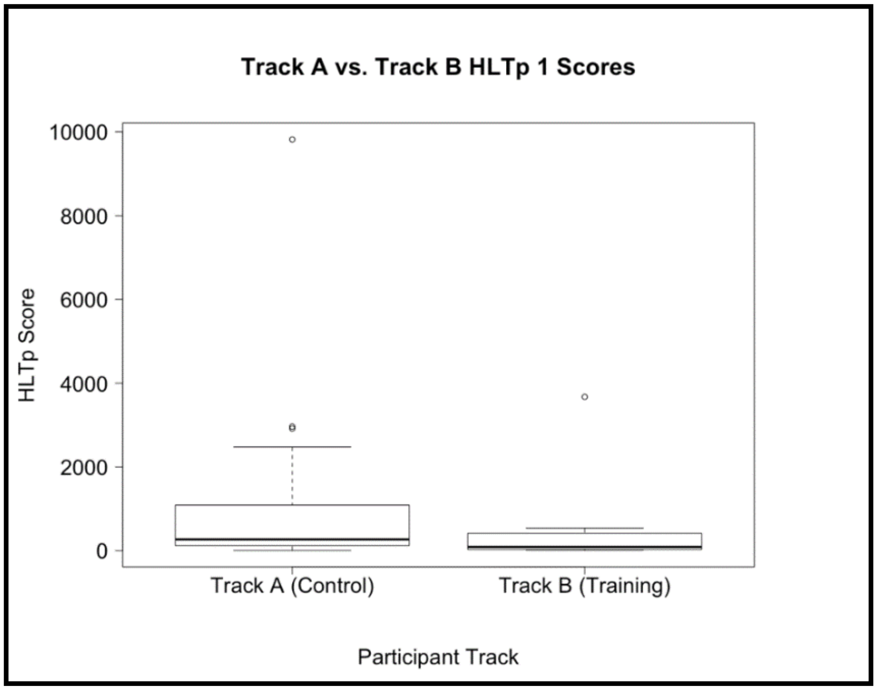


Figure 130. Track A vs. Track B HLTp 1 Scores

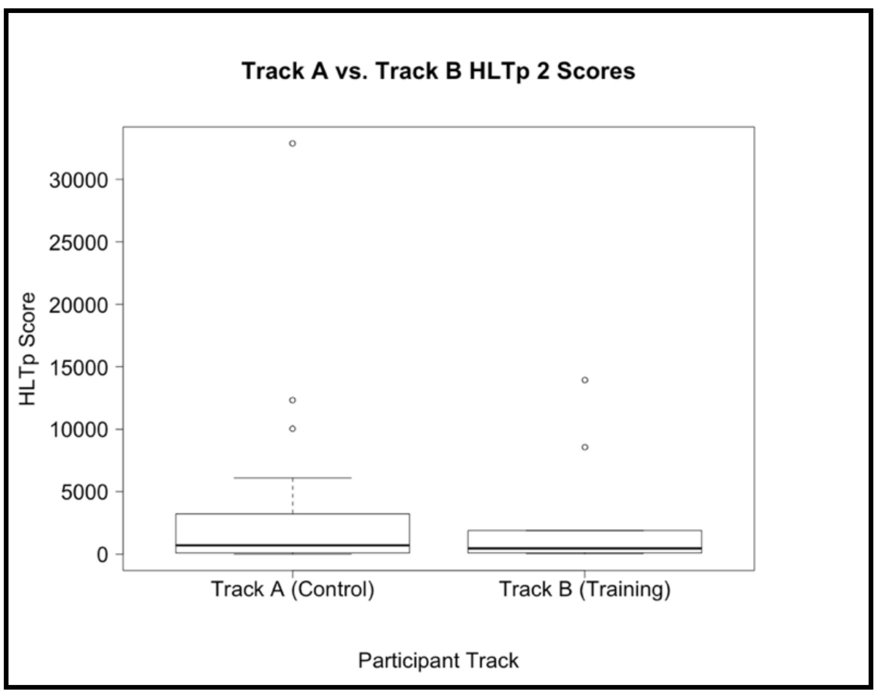


Figure 131. Track A vs. Track B HLTp 2 Scores

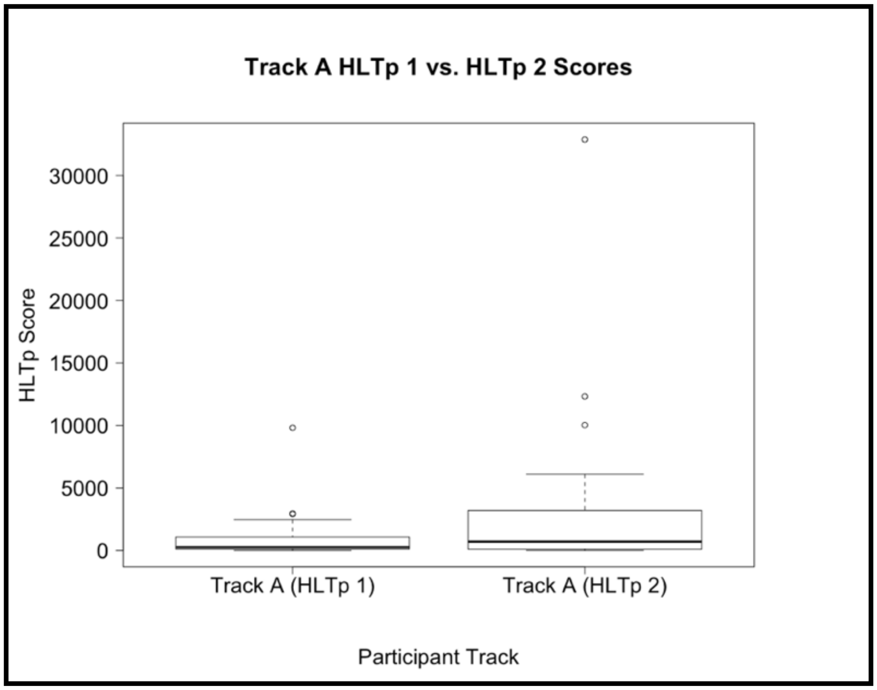


Figure 132. Track A HLTp 1 vs. HLTp 2 Scores

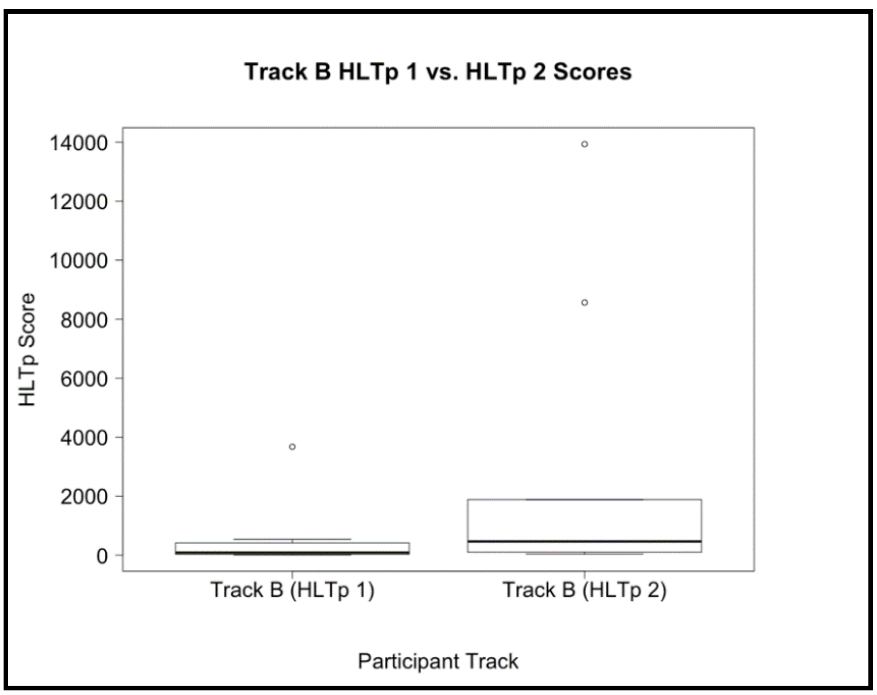


Figure 133. Track B HLTp 1 vs. HLTp 2 Scores

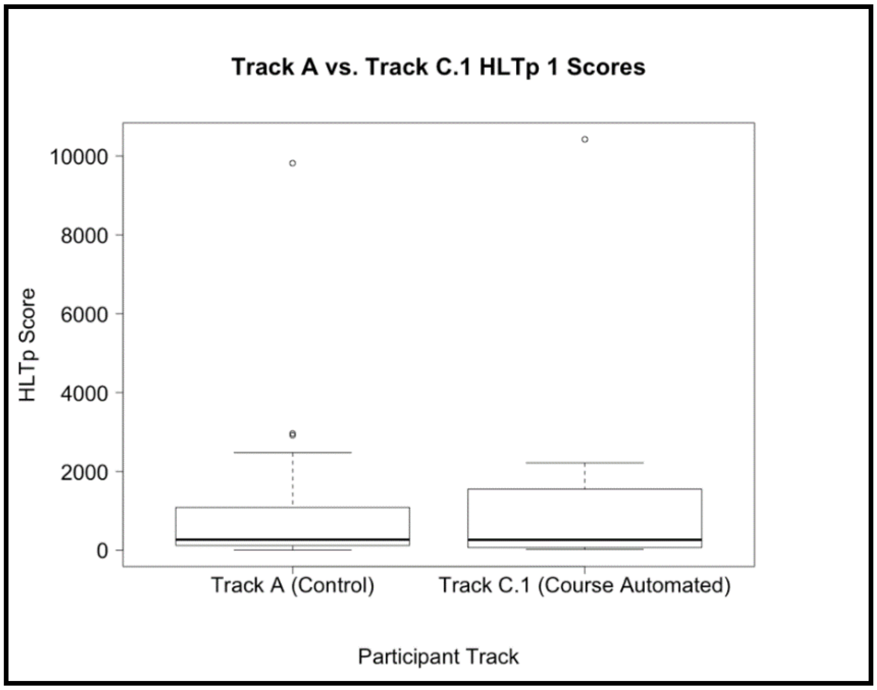


Figure 134. Track A HLTp 1 vs. Track C.1 (Course Automated) HLTp 1 Scores

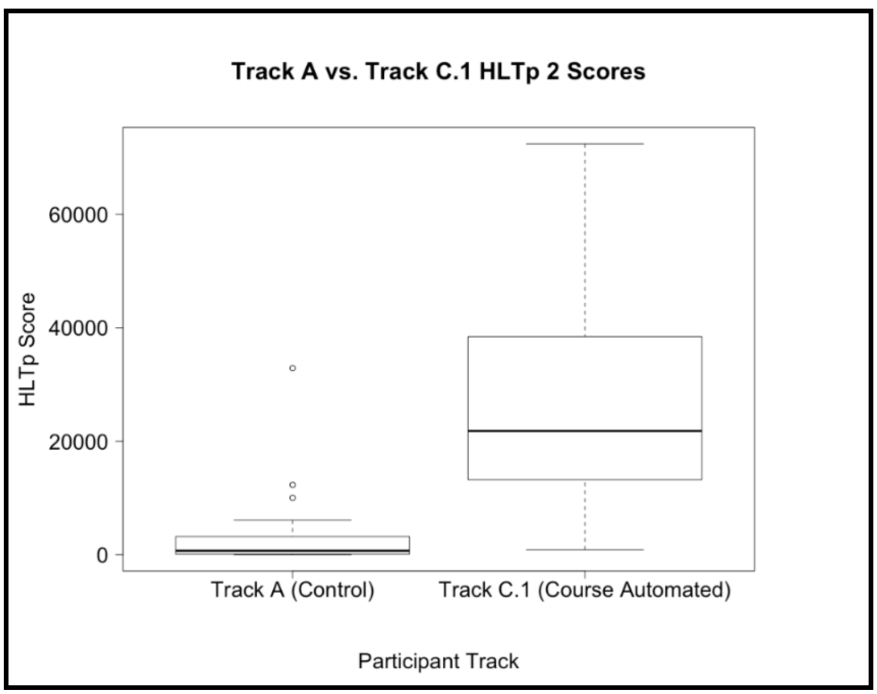


Figure 135. Track A HLTp 2 vs. Track C.1 HLTp 2 Scores

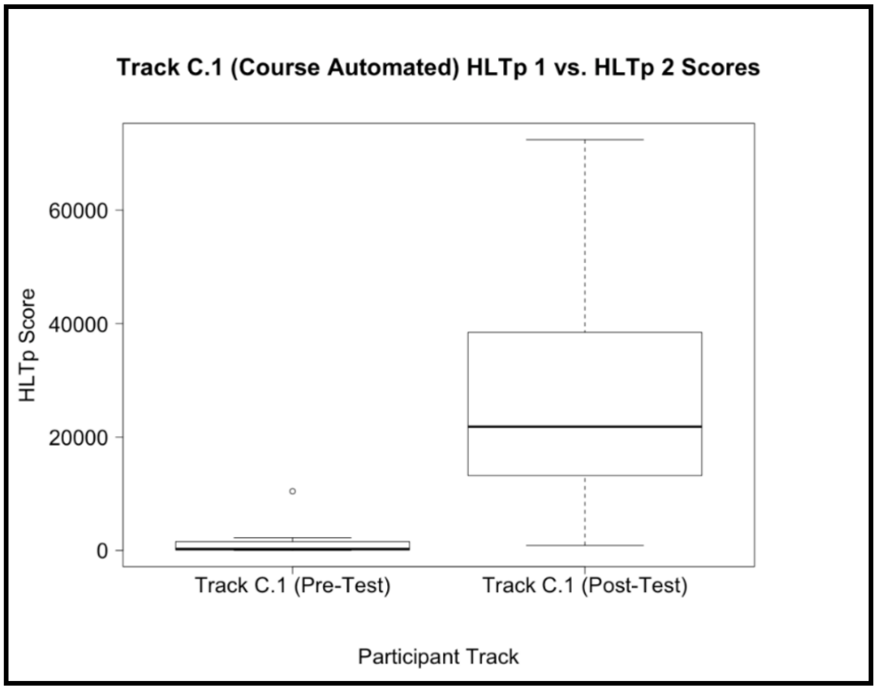


Figure 136. Track C.1 HLTp 1 vs. HLTp 2 Scores

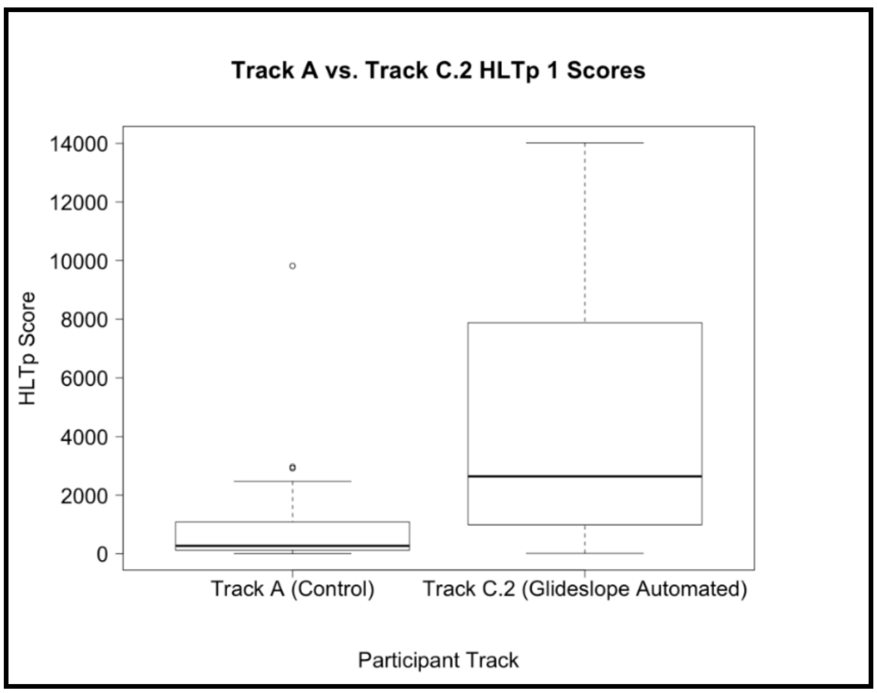


Figure 137. Track A HLTp 1 vs. Track C.2 HLTp 1 Scores

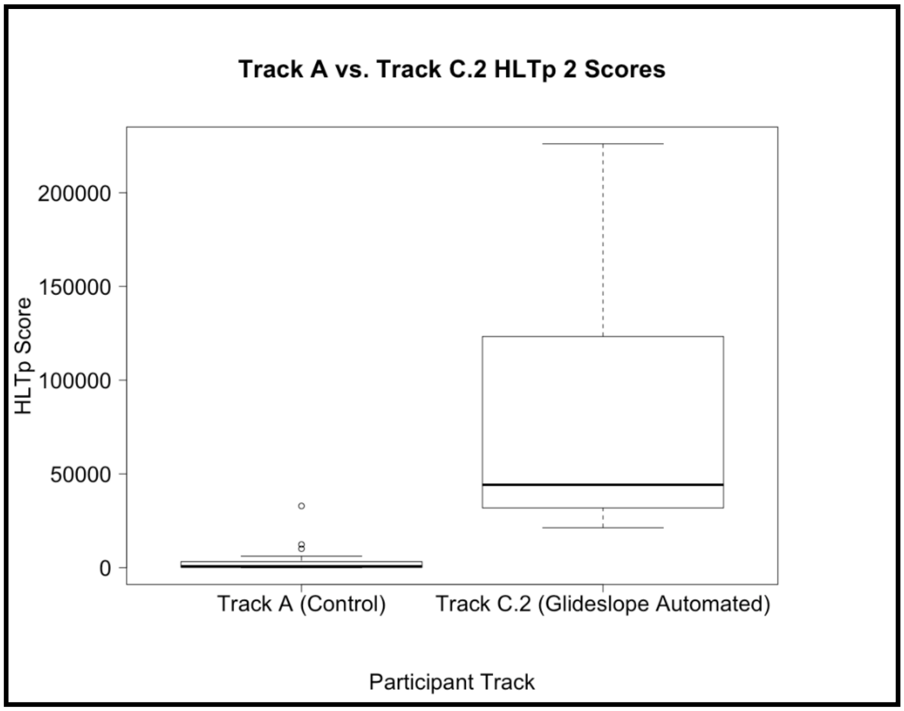


Figure 138. Track A HLTp 2 vs. Track C.2 HLTp 2 Scores

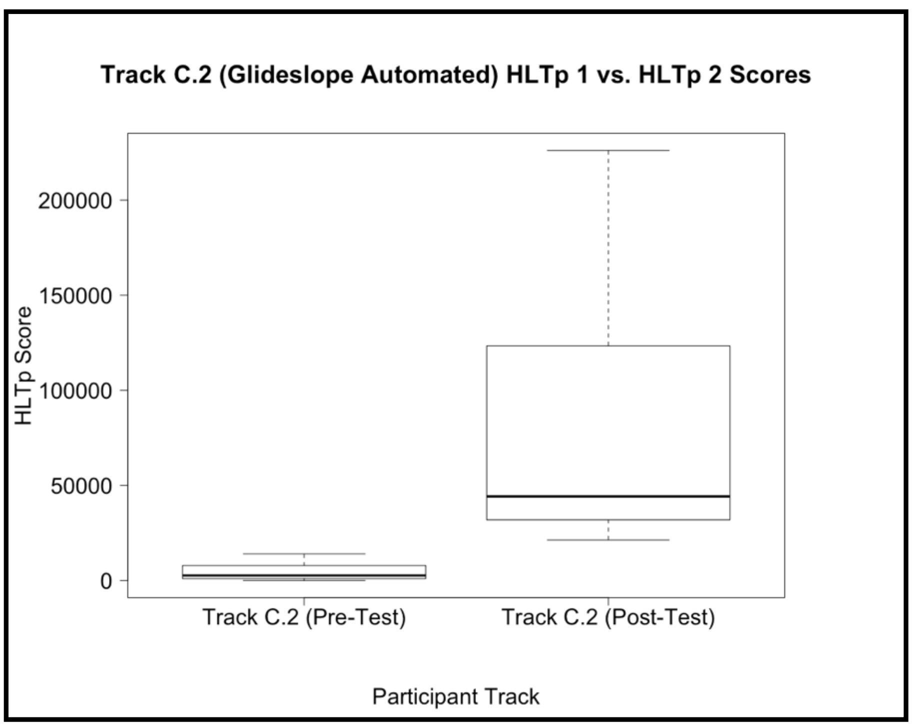


Figure 139. Track C.2 HLTp 1 vs. HLTp 2 Scores

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Balakrishnan, N., Render, B., Stair, R., & Munson, C. (2017). Frontmatter. In *Managerial Decision Modeling* (pp. i-vi). De Gruyter. https://doi.org/10.1515/9781501506208-fm

Bartone, P. T., Eid, J., Helge Johnsen, B., Christian Laberg, J., & Snook, S. A. (2009). Big five personality factors, hardiness, and social judgment as predictors of leader performance. *Leadership & Organization Development Journal*, *30*(6), 498–521. https://doi.org/10.1108/01437730910981908

Booher, H. R. (2003). *Handbook of Human Systems Integration*. John Wiley & Sons.

Boyd, J. R. (1976). *Destruction and creation*. https://www.goalsys.com/books/documents/DESTRUCTION_AND_CREATION.pdf

Buede, D. M. (2011). *The Engineering Design of Systems: Models and Methods*. John Wiley & Sons.

Casey, S. (1993). *Set phasers on stun: And other true tales of design, technology, and human errorSet Phasers on Stun*. Aegean.

Creswell, J. W., & Creswell, J. D. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.

Dantzig, G. B., & Thapa, M. N. (2006). *Linear Programming 1: Introduction*. Springer Science & Business Media.

DAU. (n.d.-a). *Adaptive Acquisition Framework*. DAU. Retrieved March 1, 2020, from https://aaf.dau.edu/

DAU. (n.d.-b). *Defense Acquisition Guide*. DAU. Retrieved March 3, 2020, from https://www.dau.edu/tools/dag/Pages/DAG-Page-Viewer.aspx?source=https://www.dau.edu/guidebooks/Shared%20Documents%20HTML/Chapter%205%20Manpower%20Planning%20and%20Human%20Systems%20Integration.aspx

Defense Safety Oversight Council. (2010). *Strategic Frameworks for Human Performance and Human Systems Integration: Spotlight on General Systems Performance Theory and Nonlinear Causal Resource Analysis.* Department of Defense.

Dell, R. (2018). *Linear Programming*. Linear Programming Lecture, Naval Postgraduate School.

Department of Defense. (2003). *Defense Acquisition System* (DoDI 5000.01). http://acqnotes.com/wp-content/uploads/2014/09/DOD-Directive-5000.01-Defense-Acquisition-System-31-Aug-2018.pdf

Department of Defense. (2018). *Summary of the 2018 National Defense Strategy*. Department of Defense. https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf

Department of Defense. (2020a). *Human systems integration (HSI) in defense acquisition* (DoDI 5000.PR).

Department of Defense. (2020b). *Operation of the Adaptive Acquisition Framework*. https://www.dau.edu/cop/risk/_layouts/15/WopiFrame.aspx?sourcedoc=/cop/risk/DAU%20Sponsored%20Documents/500002%20dodi%202015%20change1and2.pdf&action=default&DefaultItemOpen=1

Federal Aviation Administration. (2018a). *Instrument rating—airplane*. https://www.faa.gov/training_testing/testing/acs/media/instrument_rating_acs.pdf

Federal Aviation Administration. (2018b). *FAR/AIM 2019: Federal Aviation Regulations / Aeronautical Information Manual* (2019 edition). Aviation Supplies & Academics, Inc.

Fischer, C. A., Kondraske, G. V., & Stewart, R. M. (2002). Prediction of driving performance using Nonlinear Causal Resource Analysis. *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] [Engineering in Medicine and Biology*, *3*, 2473–2474. https://doi.org/10.1109/IEMBS.2002.1053382

Fleishman, E. A. (1992). *Handbook of human abilities: Definitions, measurements, and job task requirements*. Consulting Psychologists Press.

Garamone, J. (2019, Ocgtober). *Transformational change comes to DOD Acquisition Policy*. Department of Defense. https://www.defense.gov/Explore/News/Article/Article/1994041/transformational-change-comes-to-dod-acquisition-policy/

Gordon, S. (1982). Why Did Marshall Transpose the Axes? *Eastern Economic Journal*, *8*(1), 31–45. JSTOR. https://www.jstor.org/stable/40324797

Hammond, G. (2012). *The Mind of War: John Boyd and American Security*. Smithsonian Institution.

Hitchins, D. K. (1992). *Putting Systems to Work*. John Wiley & Sons.

Hughes, T. P. (2000). *Rescuing Prometheus*. Vintage Books.

iLearn2Fly. (n.d.-a). *Jeppesen Instrument Commercial DVD2* [Video]. YouTube. Retrieved March 13, 2020, from https://www.youtube.com/watch?v=iEqBeoEz-v0

iLearn2Fly. (n.d.-b). *Jeppesen Private Pilot DVD1* [Video]. YouTube. Retrieved March 13, 2020, from https://www.youtube.com/watch?v=gBXGZeYX1WQ

INCOSE. (n.d.). *Systems Engineering*. Retrieved February 22, 2020, from https://www.incose.org/systems-engineering

INCOSE. (2007). *Systems Engineering Vision 2020* (INCOSE-TP-2004-004-02, version 2.03). https://sdincose.org/wp-content/uploads/2011/12/SEVision2020_20071003_v2_03.pdf

INCOSE. (2014). *A World In Motion: Systems Engineering Vision 2025*. John Wiley & Sons.

INCOSE. (2015). *INCOSE Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities*. John Wiley & Sons. http://ebookcentral.proquest.com/lib/ebook-nps/detail.action?docID=4040424

Innes, C. R. H., Jones, R. D., Dalrymple-Alford, J. C., Hayes, S., Hollobon, S., Severinsen, J., Smith, G., Nicholls, A., & Anderson, T. J. (2007). Sensory-motor and cognitive tests predict driving ability of persons with brain disorders. *Journal of the Neurological Sciences*, *260*(1), 188–198. https://doi.org/10.1016/j.jns.2007.04.052

Jeppesen. (2016). *Jeppesen Instrument Commercial*. Jeppesen.

Kondraske, G. (2011). General systems performance theory and its application to understanding complex system performance. In *Information Knowledge Systems Management* (10th ed., pp. 235–259). IOS Press.

Kondraske, G. V. (1998). *Sample distribution: factors and influence on the quality of nonlinear causal resource analysis (NCRA) task analysis models* (Technical TR 98–005R; p. 16). The University of Texas at Arlington.

Kondraske, George V. (2019). *General Systems Performance Theory: Annotated Bibliography - External Sources* (Technical No. 2019–001R). University of Texas at Arlington.

Leibniz, G. W. F. von, Leibniz, G. W., & Montgomery, G. R. (2005). *Discourse on metaphysics and the monadology*. Courier Corporation.

Lipsey, R. G., & Chrystal, K. A. (2007). *Economics* (11th ed.). Oxford University Press.

Machol, R. E., & Goody, H. H. (1957). *System engineering: An introduction to the design of large-scale systems*. McGraw-Hill. https://nps.primo.exlibrisgroup.com/discovery/openurl?institution=01NPS_INST&vid=01NPS_INST:01NPS&sid=google&auinit=HH&aulast=Goode&title=System%20engineering:%20an%20introduction%20to%20the%20design%20of%20large-scale%20systems&genre=book&date=1957

Madni, A. M., & Orellana, D. W. (2018). Extending model-based systems engineering to address human-systems integration considerations in the system life cycle. *2018 Annual IEEE International Systems Conference (SysCon)*, 1–7. https://doi.org/10.1109/SYSCON.2018.8369498

Matsumoto, E. D., Kondraske, G. V., Ogan, K., Jacomides, L., Wilhelm, D. M., Pearle, M. S., & Cadeddu, J. A. (2006). Assessment of basic human performance resources predicts performance of ureteroscopy. *The American Journal of Surgery, 191*(6), 817–820. http://dx.doi.org/10.1016/j.amjsurg.2005.07.043

Mueller, M. W. (2008). Human systems integration—What's it all about? *INSIGHT*, *11*(2), 7–10. https://doi.org/10.1002/inst.20081127

O'Connor, P., Walliser, J., & Philips, E. (2010, October). *Evaluation of a Human Factors Analysis and Classification System Used by Trained Raters*. https://www.ingentaconnect.com/content/asma/asem/2010/00000081/00000010/art00008

Office of the Under Secretary of Defense for Acquisition & Sustainment. (n.d.). *Initiatives*. DOD Systems Engineering. Retrieved January 29, 2019, from https://www.acq.osd.mil/se/initiatives/init_hsi.html

O*NET. (n.d.). *Summary report for: 53–2012.00—Commercial pilots*. O*NET OnLine. Retrieved January 2, 2020, from https://www.onetonline.org/link/summary/53-2012.00

Zimmerman, P. (2014, April 3). *A review of Model-Based Systems Engineering practices and recommendations for future Directions in the Department of Defense*. 2nd Systems Engineering in the Washington Metropolitan Area (SEDC 2014) Conference, Chantilly, VA.

Pidd, M. (2009). *Tools for thinking: Modelling in management science*. John Wiley & Sons.

Proctor, R. W., & Van Zandt, T. (2008). *Human Factors in simple and complex systems.* 2nd ed. CRC Press.

R Core Team, (2019). *R: A language and environment for statistical computing.* https://www.R-project.org/

Rardin, R. L. (2016). *Optimization in operations research*. Pearson Education.

Shattuck, L. G. (2017, October 25). *From cognitive systems engineering to human systems integration: A short but necessary journey*. CRC Press. https://doi.org/10.1201/9781315572529-17

Shishko, R. G. et al. (1992). *NASA systems engineering handbook*. https://ntrs.nasa.gov/search.jsp?R=19930011999

Simon, H. A. (1955). Behavioral model of rational choice. *Quarterly Journal of Economics*. https://academic.oup.com/qje/article/69/1/99/1919737

Tvaryanas, A. (2010). *A Discourse in human systems integration* [Dissertation, Naval Postgraduate School]. https://calhoun.nps.edu/handle/10945/10571

U.S. Air Force. (2012). *Report on* Aircraft Oxygen Genera*tion*. U.S. Air Force. https://www.airforcemag.com/PDF/DocumentFile/Documents/2012/AFSAB_Oxygen_020112.pdf

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California